Human Inference

Klant	Ministerie van Volksgezondheid, Welzijn en Sport
Project	LRZa
Onderdeel	Analyse en Advies



Document informatie

Klant	Ministerie van Volksgezondheid, Welzijn en Sport
Project	LRZa
Onderdeel	Analyse en Advies
Versie	Definitief
Bestand	2014-12-04-LRZa_Analyse_en_Advies_def.docx
Creatiedatum	3-9-2014
Datum wijziging	04-12-2014
Datum afdruk	04-12-2014
Auteur	Mark Ansink / Ron Mulderij / Dennis Groenhuijzen
Status	Definitief

Distributie

Bedrijf	Contactpersoon
MinVWS	[redacted]
MinVWS	[redacted]
Vektis	[redacted]
Handelsregister	[redacted]
Handelsregister	[redacted]
BIG	[redacted]
BIG	[redacted]
Human Inference	Dennis Groenhuijzen
Human Inference	Mark Ansink
Human Inference	Ron Mulderij

Inhoudsopgave

Bijlagen

1 Inleiding

Dit document beschrijft de resultaten van de analyse ten behoeve van het Landelijk Register Zorgaanbieders (LRZa), zoals deze is uitgevoerd door Human Inference.

Het ministerie van Volksgezondheid, Welzijn en Sport werkt aan een landelijk register voor de zorgsector, naar aanleiding van de wet kwaliteit, klachten en geschillen zorg (Wkkgz). Het register ontsluit verschillende bronregistraties met gegevens over zorgondernemingen en zorgverleners en de vestigingen (zie definities in bijlage 6) waar zorg wordt geleverd Bronnen voor het LRZa zijn onder andere het Handelsregister (KvK), het AGB-register (Vektis) en het BIG-register (CIBG).

Human Inference is een bedrijf dat intelligente softwarecomponenten ontwikkelt die bedrijven en organisaties ondersteunen bij het verbeteren van datakwaliteit van relatiegegevens. Hierdoor is het mogelijk betere service te verlenen, risico's beter te managen en meer cross- en upsell te realiseren. Human Inference heeft veel ervaring op het gebied van analyse van data en met de implementatie en integratie van deze software componenten in complexe ICT-omgevingen (zie bijlage 5 voor een nadere beschrijving).

Het ministerie van Volksgezondheid, Welzijn en Sport heeft Human Inference gevraagd een analyse uit te voeren en een advies uit te brengen met betrekking tot de bronnen die voor de vulling van het LRZa gebruikt gaan worden.

1.1 Doelstelling

Gerekend naar koppelbaarheid is de kwaliteit van de gegevens in de bronregistraties naar verwachting niet optimaal.

Inzicht is gewenst in:

- de matchingsgraad, ofwel de mate waarin de gegevens vanuit de verschillende bronnen daadwerkelijk eenduidig met elkaar gematcht kunnen worden;
- de kwaliteit van de gegevens binnen de bronnen zelf, gericht op zaken als vullingsgraad, eenduidige wijze van registreren, actualiteit, etc.

Het uit onderzoek verkregen inzicht moet de basis vormen voor een concrete aanpak tot initiële en structurele verbetering van de gegevenskwaliteit in de bronregistraties, teneinde te komen tot een juist en volledig LRZa.

Concreet levert dit onderzoeksproject het inzicht in de volgende punten:

- Vullingsgraad van de afzonderlijke datasets;
- Correcte vulling van de rubrieken;
- Juistheid van de gegevens;
- Voorkomende fictieve waarden;
- Mogelijke geautomatiseerde verbeteringen;
- Mate van relateerbaarheid van de gegevens over de verschillende bronnen.

1.2 Opdrachtverstrekking

De opdracht voor het uitvoeren van deze analyse is op 27 juni 2014 door het Ministerie van Volksgezondheid, Welzijn en Sport aan Human Inference verstrekt. In deze opdracht zijn de volgende op te leveren producten benoemd:

1. Analyse van afzonderlijke datasets van de drie data-leveranciers van de LRZa en een matching van deze drie bronnen, waardoor een beeld ontstaat van aantallen zekere matches, aantallen mogelijke matches en aantallen niet te matchen records;
2. Een nader onderzoek welke verbetermogelijkheden kunnen worden toegepast;
3. Een advies m.b.t. de vervolgstappen om te komen tot een operationeel kwalitatief geborgd LRZa;
4. Een schriftelijke rapportage over bovenstaande producten.

1.3 Structuur van dit rapport

Dit rapport geeft een samenvatting van de resultaten van de uitgevoerde analyse. Hierbij is de volgende indelingen aangehouden:

- Hoofdstuk 2 beschrijft de onderzoeksmethode, zoals die ook is afgestemd bij de opdrachtverstrekking;
- Hoofdstuk 3 bevat informatie over de aangeleverde datasets;
- Hoofdstuk 4 bevat de bevindingen ten aanzien van de kwaliteit van de aangeleverde datasets ten aanzien van vullingsgraad en inhoudelijke kwaliteit;
- Hoofdstuk 5 beschrijft de beoogde samenhang van de datasets;
- Hoofdstuk 6 geeft inzicht in de werkelijke samenhang van de datasets op basis van een uitgevoerde matching;
- Hoofdstuk 7 geeft inzicht in mogelijkheden voor verrijking;
- Hoofdstuk 8 bevat tot slot de conclusies en aanbevelingen.

Daarnaast is een aantal bijlagen opgenomen:

- Bijlage 1 geeft een overzicht van de activiteitcodes van AGB;
- Bijlage 2 geeft een overzicht van SBI-codes volgens het Handelsregister;
- Bijlage 3 geeft een overzicht van de beroepsgroepen volgens BIG;
- Bijlage 4 geeft gedetailleerde informatie over de uitgevoerde datakwaliteitsanalyse;
- Bijlage 5 geeft een beschrijving van Human Inference;
- Bijlage 6 bevat definities van onder meer de begrippen "zorgonderneming", "zorgverlener" en "zorgvestiging".

2 Onderzoeksmethode

Conform de opdrachtverstrekking zijn voor de uitvoering de volgende stappen doorlopen.

Stap 1

Aanlevering van de datasets van AGB, BIG en Handelsregister vanuit de bronhouder aan Human Inference.

Laden van deze datasets door Human Inference ten behoeve van de analyse.

Stap 2

Analyse van de afzonderlijke datasets, waarbij onder meer is gelet op:

- Vullingsgraad;
- Correcte vulling van rubrieken (waaronder validatie van NAW-gegevens);
- Fictieve waarden;
- Mogelijke geautomatiseerde verbeteringen;
- Mogelijk onjuiste KvK-nummers i.r.t. de NAW-gegevens.

Matching van de datasets zodat een beeld ontstaat van aantallen zekere matches, aantallen mogelijke matches en aantallen niet te matchen records.

Stap 3

Na het uitvoeren van de analyses van stap 2 is onderzocht welke verbetermogelijkheden toegepast kunnen worden, zodat tot een betere vulling van het LRZa gekomen kan worden.

Stap 4

Presentatie van de resultaten en oplevering van het rapport en advies.

3 Datasets

In dit hoofdstuk zullen de datasets beschreven worden zoals deze zijn aangeleverd door de verschillende partijen. In bijlage 2,3 en 4 zijn per bron en type de aantallen per SBI-code of Zorgsoort of Beroepsgroep opgenomen.
In de volgende hoofdstukken wordt gesproken over zorgondernemingen, zorgvestigingen en zorgverleners. In bijlage 6 staan de exacte definities beschreven.

3.1 Vektis

Vektis heeft vanuit AGB de volgende entiteiten aangeleverd, zoals hierboven al te zien is:

- Zorgverleners;
- Zorgondernemingen;
- Zorgvestigingen.

Deze gegevens zijn aangeleverd in XML files die zijn opgebouwd volgens het datamodel van Vektis. Concreet betekende dit dat de gegevens van de zorgverleners, zorgondernemingen en zorgvestigingen verspreid staan over meerdere bestanden.
Ten behoeve van het onderzoek hebben wij met behulp van de beschrijving die is meegeleverd de bestanden aan elkaar gekoppeld.
Op basis hiervan zijn de volgende aantallen per type verkregen:

Zorgondernemingen	107.949
Zorgvestigingen	115.725
Zorgverleners	163.065

3.2 Handelsregister

Vanuit het Handelsregister is één bestand aangeleverd waarin zowel de zorgondernemingen als de vestigingen vermeld staan. Het gaat daarbij om een extractie op basis van zorg gerelateerde SBI-codes voor de hoofd- dan wel nevenactiviteit.
De geleverde dataset bevatte 241.413 ondernemingen/vestigingen. Hiervan hebben 163.957 records een hoofdvestigingsindicatie, 36.625 records hebben een nevenvestigingsindicatie en de overige 40.831 records hebben geen indicatie (registraties zonder feitelijke vestiging).

3.3 CIBG

Het CIBG heeft vanuit het BIG-register een bestand met in totaal 365.385 registraties aangeleverd.

4 Analyse datasets

Bij de analyse van de datasets is primair gelet op de gegevens die nodig zijn om het LRZa correct te kunnen realiseren. Aandachtspunten daarbij zijn onder meer:

- Vullingsgraad;
- Fictieve waarden;
- Kwaliteit van tenaamstellingen;
- Kwaliteit van adresgegevens;
- Kwaliteit van telefoonnummers;
- Kwaliteit van e-mail adressen.

De vullingsgraad van de aangeleverde bestanden is van belang om tot een voldoende hoge mate van matching te kunnen komen. Immers, hoe meer gegevens bekend zijn, hoe beter een oordeel gegeven kan worden.

Voor een optimale kwaliteit is het van belang fictieve waarden te signaleren, zodat daarmee in de opbouw rekening gehouden kan worden. Denk aan zaken als een standaard geboortedatum "01-01-1900" indien deze onbekend is of "Onbekend" in plaats van een veld leeg te laten. Ook de kwaliteit van de matching neemt hiermee toe.

Kwaliteit van tenaamstellingen, adressen, telefoonnummers en e-mail adressen is van belang om enerzijds de juiste gegevens te registreren, maar ook om gegevens in een consistent formaat te registreren. Ook de kwaliteit van de matching neemt hiermee toe. Voor het bepalen van de kwaliteit van tenaamstellingen, adressen, telefoonnummers en e-mail adressen is gebruik gemaakt van software van Human Inference. Deze software maakt gebruik van fouttolerante technieken en kennisdatabases om een optimale interpretatie en beoordeling te realiseren (zie ook bijlage 5 "Over Human Inference").

4.1 AGB-register

Een algemene opmerking ten aanzien van het AGB-register is dat de tenaamstellingen in alle gevallen volledig in hoofdletters zijn opgenomen. Deze tenaamstellingen kunnen, ten behoeve van de presentatie van de gegevens in het LRZa, geautomatiseerd worden omgezet naar een combinatie van hoofdletters en kleine letters. Er zullen hierbij geen automatische correcties worden doorgevoerd in de bronnen.

4.1.1 AGB Zorgondernemingen

Het aantal aangeleverde zorgondernemingen bedraagt 107.949. Een verdeling hiervan naar de verschillende zorgsoorten is opgenomen in bijlage 1.

De vullingsgraad van de gegevens die voor de matching en de opbouw van het LRZa nodig zijn is hieronder weergegeven (inclusief een voorbeeld van een mogelijke inhoud):

Veld	**Voorbeeld**	**Vullingsgraad**
handelsnaam1	PRAKTIJK VOOR HUID- EN OEDEEMTHERAP	100%
hrnummer		3,9%
Adresgegevens:		
straatnaam	[redacted]	100%
huisnummer	[redacted]	100%
huisnummertoevoeging		9,9%
postcode	[redacted]	100%
plaats	[redacted]	100%
Communicatiegegevens:		
e-mail adres	[redacted]	22,7%
telefoonnummer	[redacted]	81,4%
fax	[redacted]	4,1%

Het meest opvallende is de lage vullingsgraad van de het handelsregisternummer. In slechts 4.262 gevallen (3,9%) is dat gevuld. Het handelsregisternummer is een essentieel element bij het relateren van de AGB-gegevens aan de gegevens in het Handelsregister. Hierbij moet nog worden aangetekend dat het bovendien gaat om veel dezelfde handelsregisternummers, want het aantal unieke nummers is slechts 2.547 (59,8%). Een voorbeeld betreft handelsregisternummer 27310787 van stichting Middin dat bijvoorbeeld 43 keer voorkomt bij verschillende vestigingen. Het vestigingsnummer is derhalve ook een essentieel element voor het leggen van de juiste relaties.

Vanuit de kwaliteitsanalyse van de gegevens kan worden geconcludeerd dat de gegevens redelijk tot goed zijn. Er zijn echter wel een aantal bevindingen die aandacht behoeven om tot een kwalitatief goede LRZa te komen.
De volgende resultaten zijn te melden:

- Over de inhoudelijke juistheid van de tenaamstelling is geen oordeel mogelijk; dit kan wel op basis van matching resultaten en in vergelijking met het Handelsregister;

- Alle tenaamstelling zijn in hoofdletters opgenomen. Voor vergelijking met de tenaamstelling in het Handelsregister is het nodig de tenaamstelling in AGB geautomatiseerd om te zetten naar de schrijfwijze van het Handelsregister. Deze functionaliteit is aan LRZa toe te voegen en behoeft niet te leiden tot aanpassing van de schrijfwijze in het AGB-register;
- Bij 1,2% van de adressen is een geautomatiseerde correctie/verbetering mogelijk. Er zullen hierbij geen automatische correcties worden doorgevoerd in de bronnen;
- Bij 0,2% van de adressen zijn de gegevens onbetrouwbaar of onjuist en niet geautomatiseerd te verbeteren; handmatige correctie in het bronregister is hiervoor nodig;
- Bij 0,3% van de telefoonnummers zijn de gegevens onbetrouwbaar of onjuist; handmatige correctie in het bronregister is hiervoor nodig.

Details van de kwaliteitsanalyse zijn opgenomen in bijlage 4.
Ongecorrigeerd opnemen van deze gegevens in het LRZa kan leiden tot onjuiste of incomplete informatievoorziening bij gebruik van het LRZa.

4.1.2 AGB Zorgvestigingen

In totaal zijn er 115.725 zorgvestigingen aangeleverd. Hiervan hebben 61 een handelsregister vestigingsnummer gevuld. Dit nummer is een essentieel element bij het leggen van relaties tussen AGB-register en Handelsregister. Een verdeling van de zorgvestigingen naar de verschillende zorgsoorten is opgenomen in bijlage 1.

De vullingsgraad van de gegevens die voor de matching en de opbouw van het LRZa nodig zijn is hieronder weergegeven (inclusief een voorbeeld van een mogelijke inhoud):

Veld	**Voorbeeld**	**Vullingsgraad**
<u>Naamgegevens:</u>		
Handelsnaam1	PIETERNEL MAMMACARE SPECIALIST	100%
vestigingsnummer		0%
<u>Adresgegevens:</u>		
straatnaam	[redacted]	100%
huisnummer	[redacted]	100%
huisnummertoevoeging		9,9%
postcode	[redacted]	100%
plaats	[redacted]	100%
<u>Communicatiegegevens:</u>		
e-mail adres	[redacted]	20,9%
telefoonnummer	[redacted]	76,6%

Vanuit de kwaliteitsanalyse van de gegevens kan worden geconcludeerd dat de gegevens redelijk tot goed zijn. Er zijn echter wel een aantal bevindingen die aandacht behoeven om tot een kwalitatief goede LRZa te komen.
De volgende resultaten zijn te melden:

- Over de inhoudelijke juistheid van de tenaamstelling is geen oordeel mogelijk; dit kan wel op basis van matching resultaten en in vergelijking met het Handelsregister;
- Bij 1,2% van de adressen is een geautomatiseerde correctie/verbetering mogelijk. Er zullen hierbij geen automatische correcties worden doorgevoerd in de bronnen;
- Bij 0,2% van de adressen zijn de gegevens onbetrouwbaar of onjuist en niet geautomatiseerd te verbeteren; handmatige correctie in het bronregister is hiervoor nodig;
- Bij 0,3% van de telefoonnummers zijn de gegevens onbetrouwbaar of onjuist; handmatige correctie in het bronregister is hiervoor nodig.

Details van de kwaliteitsanalyse zijn opgenomen in bijlage 4.
Ongecorrigeerd opnemen van deze gegevens in het LRZa kan leiden tot onjuiste of incomplete informatievoorziening bij gebruik van het LRZa.

4.1.3 AGB Zorgverleners

Het aantal aangeleverde zorgverleners bedraagt 163.065. Een verdeling hiervan naar de verschillende zorgsoorten is opgenomen in bijlage 1.
Vanuit AGB gezien zouden er 80.428 zorgverleners bekend moeten zijn in het BIG-register aangezien deze voldoen aan de zorgsoorten die een BIG-registratie moeten hebben en waarbij de AGB-code geen einddatum heeft.

De vullingsgraad van de gegevens die voor de matching en de opbouw van het LRZa nodig zijn is hieronder weergegeven (inclusief een voorbeeld van een mogelijke inhoud):

Veld	Voorbeeld	Vullingsgraad
Naamgegevens:		
voorletters	EPM	100%
voorvoegsel		23%
achternaam	BOLSIUS	100%
geboortenaamvoorvoegsel		0%
geboortenaam		0%
datumgeboorte	[redacted]	100%
geslacht	[redacted]	100%
Adresgegevens:		
straatnaam	[redacted]	100%
huisnummer	[redacted]	100%
huisnummertoevoeging	[redacted]	12,2%
postcode	[redacted]	100%
plaats	[redacted]	100%
Communicatiegegevens:		
telefoonnummer	[redacted]	94,1%
e-mail adres	[redacted]	72,7%
Referentie (BIG nummer)		99,5%

De vulling van het BIG-nummer bedraagt ten opzichte van de set die een BIG-registratie moet hebben 99.5%. Hierdoor is in 0.5 % van de gevallen (402 registraties) geen koppeling op basis van dit nummer mogelijk. Daarnaast zijn er 297 gevallen waarbij het registratienummer gevuld is met een fictieve waarde ("0000000000"). Ook zijn er 476 BIG-nummers in de set die niet bekend zijn in het BIG-register. Enkele voorbeelden zijn hieronder opgenomen:

BIG-nummer	Naam	Straat/huisnr	Postcode/plaats
39034906104	JA SANDERS	[redacted]	[redacted]
09034488404	FPJM MICHIELSEN	[redacted]	[redacted]
09913173302	I TALHAOUI	[redacted]	[redacted]

Vanuit de kwaliteitsanalyse van de gegevens kan worden geconcludeerd dat de gegevens redelijk tot goed zijn. Er zijn echter wel een aantal bevindingen die aandacht behoeven om tot een kwalitatief goed LRZa te komen.
De volgende resultaten zijn te melden:

- Bij 2,9% van de tenaamstellingen zijn onjuistheden geconstateerd; handmatige correctie in het bronregister is hiervoor nodig;
- Bij 1,2% van de adressen is een geautomatiseerde correctie/verbetering mogelijk. Er zullen hierbij geen automatische correcties worden doorgevoerd in de bronnen;
- Bij 0,2% van de adressen zijn de gegevens onbetrouwbaar of onjuist en niet geautomatiseerd te verbeteren; handmatige correctie in het bronregister is hiervoor nodig;
- Bij 2,5% van de e-mail adressen zijn de gegevens onbetrouwbaar of onjuist; handmatige correctie in het bronregister is hiervoor nodig;
- Bij 0,3% van de telefoonnummers zijn de gegevens onbetrouwbaar of onjuist; handmatige correctie in het bronregister is hiervoor nodig;

Details van de kwaliteitsanalyse zijn opgenomen in bijlage 4.
Ongecorrigeerd opnemen van deze gegevens in het LRZa kan leiden tot onjuiste of incomplete informatievoorziening bij gebruik van het LRZa.

4.2 Handelsregister

Het handelsregister heeft 241.413 ondernemingen en rechtspersonen aangeboden. Welke onderverdeeld kunnen worden naar de SBI-codes zoals weergegeven in bijlage 2.

De vullingsgraad van de gegevens die voor de matching en de opbouw van het LRZa nodig zijn, is hieronder weergegeven (inclusief een voorbeeld van een mogelijke inhoud):

Veld	Voorbeeld	Vullingsgraad
Naamgegevens:		
handelsnaam	Drogisterij Helfrich B.V.	100%
dossiernummer	01000065	100%
Vestigingsnummer	000019421877	82,8%
Vestigingsadres:		
Straatnaam	[redacted]	99,8%
Huisnummer	[redacted]	99,9%
Toevoeging huisnummer		15%
Postcode	[redacted]	99,9%
Woonplaats_nen	[redacted]	99,9%
Correspondentieadres:		
Straatnaam	[redacted]	99,8%
Huisnummer	[redacted]	99,9%
Toevoeging huisnummer		12,2%
Postcode	[redacted]	99,6%
Woonplaats_nen	[redacted]	99,9%
Communicatiegegevens:		
Telefoonnummer	[redacted]	54%
Mobiele telefoon		42%

In de aangeleverde dataset is een indicatie “hoofdvestiging” opgenomen. Voor de matching van AGB tegen het Handelsregister zijn in alle gevallen de AGB-records, zowel zorgondernemingen als zorgvestigingen, tegen de volledige aangeleverde subset uit het Handelsregister gematcht.

De vullingsgraad van het vestigingsnummer lijkt met 82,8% laag, maar is verklaarbaar aangezien niet bij alle registraties in het Handelsregister een vestigingsnummer bekend hoeft te zijn.

Handelsnaam	Straat	Huisnr	Postcode	Plaats
Buurtvereniging Appel-Driedorp	[redacted]	[redacted]	[redacted]	[redacted]
Stichting Een Helpende Hand	[redacted]	[redacted]	[redacted]	[redacted]
Stichting Wetenschapsbevordering Klinisch psycholoog en Klinisch neuropsycholoog	[redacted]	[redacted]	[redacted]	[redacted]
Stichting 4our Miraculous Trust	[redacted]	[redacted]	[redacted]	[redacted]
Stichting Manueel Praktizijn	[redacted]	[redacted]	[redacted]	[redacted]
Stichting Ouderenwerk Hollandscheveld	[redacted]	[redacted]	[redacted]	[redacted]
Stichting Ondersteuning Ziekenhuiszorg	[redacted]	[redacted]	[redacted]	[redacted]

Voor tenaamstellingen geldt het Handelsregister als authentieke registratie, voor de persoonsgegevens hanteert het Handelsregister de Basisregistratie Personen (BRP v/h GBA) en voor adressen in Nederland hanteert het Handelsregister de Basisregistratie Adressen en Gebouwen (BAG), zodat de kwaliteit hiervan als een gegeven is beschouwd.
Er zijn echter wel een aantal bevindingen die aandacht behoeven:

- Bij 0,8% van de adressen zijn de gegevens niet eenduidig te standaardiseren. Hierbij gaat het in veel gevallen om buitenlandse addressen waarbij geen landcode bekend is; handmatige correctie in het bronregister is hiervoor nodig;
- Bij 0,6% van de telefoonnummers zijn de gegevens onbetrouwbaar of onjuist; handmatige correctie in het bronregister is hiervoor nodig.

Details van de kwaliteitsanalyse zijn opgenomen in bijlage 4.
Ongecorrigeerd opnemen van deze gegevens in het LRZa kan leiden tot onjuiste of incomplete informatievoorziening bij gebruik van het LRZa.

4.3 BIG-register

Er zijn in totaal 365.385 registraties aangeleverd, welke onderverdeeld kunnen worden in de beroepsgroepen die in bijlage 3 zijn opgenomen.

De vullingsgraad van de gegevens die voor de matching en de opbouw van het LRZa nodig zijn is hieronder weergegeven (inclusief een voorbeeld van een mogelijke inhoud):

Veld	**Voorbeeld**	**Vullingsgraad**
Voorvoegsel	van der	23,5%
Achternaam	A	100%
Voorletters	M.M.A	99,9%
Geslacht	[redacted]	100%
Voorvoegsel partnernaam		6,6%
Partnernaam	[redacted]	27,3%
Artikelregistratienummer	[redacted]	100%
Werkadres 1:		
Straat	[redacted]	58,5%
Huisnr	[redacted]	58,6%
Plaats	[redacted]	58,6%
Postcode	[redacted]	58,4%
Landcode	[redacted]	59,1%
Werkadres 2:		
Straat		3,7%
Huisnr		3,7%
Plaats		3,7%
Postcode		3,7%
Landcode		3,8%
Werkadres 3:		
Straat		0,5%
Huisnr		0,5%
Plaats		0,5%
Postcode		0,5%
Landcode		0,5%

Ten aanzien van de vullingsgraad van de gegevens in het BIG-register zijn een aantal opmerkingen te maken:

- Het BIG-register kent een koppeling met de Basisregistratie Personen, echter ten behoeve van het LRZa wordt het woonadres niet aangeleverd;
- In de aanlevering voor de vulling van het LRZa zijn alleen werkadressen opgenomen;
- Bij niet alle registraties (0,1%) zijn de voorletters ingevuld;
- Bij ruim 41% van de registraties is geen werkadres bekend;

- BIG registratie nummer is niet uniek in de aangeleverde set. Er zijn 364.702 unieke waardes (99.8%) gevonden. Uit analyse van een aantal van deze dubbele registratie nummers blijkt dat het BIG register deze als 1 persoon heeft geregistreerd maar met meerdere specialismen. In het aangeleverde bestand heeft dit geresulteerd in meerdere records van deze personen.

Het feit dat alleen werkadressen bekend zijn en woonadressen ontbreken, kan leiden tot een minder volledig en minder betrouwbaar LRZa en maakt bovendien de identificatie van de juiste zorgverlener aanmerkelijk lastiger.

Vanuit de kwaliteitsanalyse van de gegevens kan worden geconcludeerd dat de gegevens over het geheel genomen goed zijn. Voor tenaamstellingen kent het BIG-register een koppeling met de Basisregistratie Personen (BRP v/h GBA), zodat de kwaliteit hiervan als een gegeven is beschouwd.
Er zijn echter wel een aantal bevindingen die aandacht behoeven om tot een kwalitatief goed LRZa te komen:

- Bij 3,6% van de werkadressen is een geautomatiseerde correctie/verbetering mogelijk. Er zullen hierbij geen automatische correcties worden doorgevoerd in de bronnen;
- Bij 0,7% van de werkadressen zijn de gegevens onbetrouwbaar of onjuist en niet geautomatiseerd te verbeteren; handmatige correctie in het bronregister is hiervoor nodig.

Details van de kwaliteitsanalyse zijn opgenomen in bijlage 4.
Ongecorrigeerd opnemen van deze gegevens in het LRZa kan leiden tot onjuiste of incomplete informatievoorziening bij gebruik van het LRZa.

5 Beoogde samenhang datasets

In dit hoofdstuk wordt de beoogde samenhang tussen de verschillende bronnen voor de verschillende entiteiten besproken. Vertrekpunt voor de samenhang is het AGB-register, waarvoor een koppeling met het Handelsregister en het BIG-register gelegd moet worden.

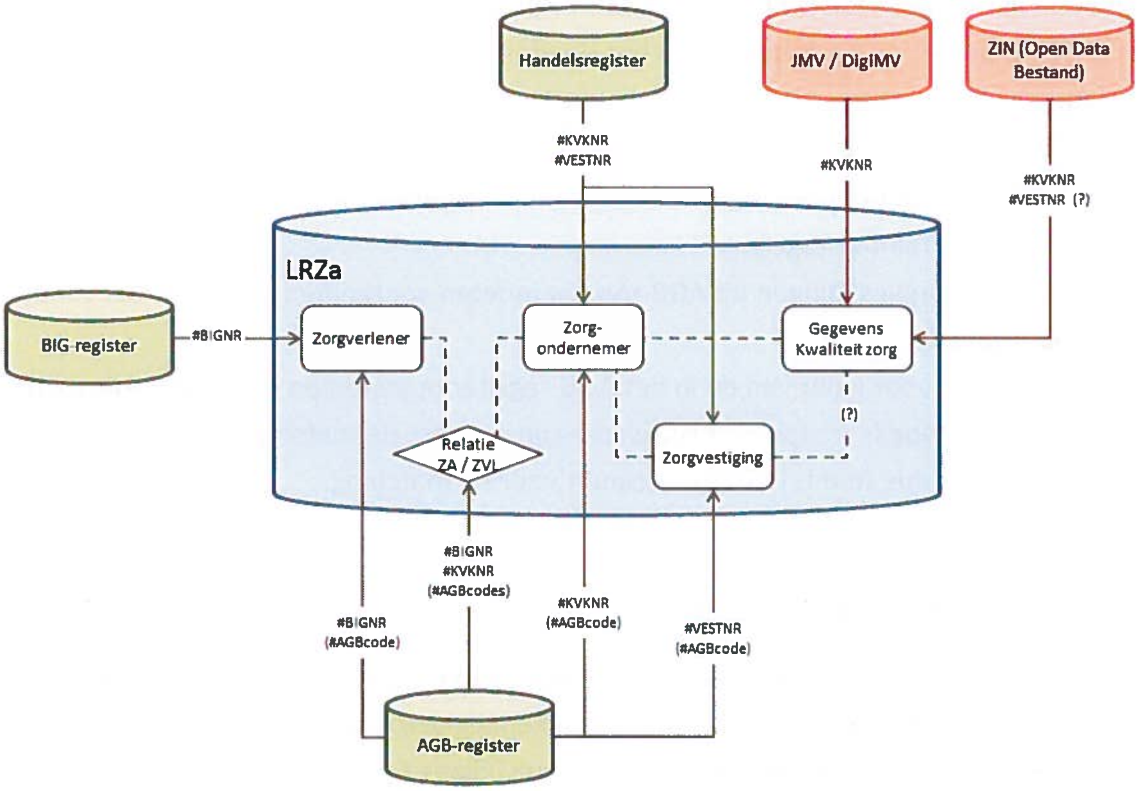
Bron	Type	Aantal	Opmerkingen
AGB	Zorgondernemingen	107.949	*Zouden moeten voorkomen in de extractie vanuit het Handelsregister*
AGB	Zorgvestigingen	115.725	*Zouden moeten voorkomen in de extractie vanuit het Handelsregister*
AGB	Zorgverleners	163.065	*Zouden, m.u.v. paramedici, moeten voorkomen in BIG*
Handelsregister (HR)	Ondernemingen	241.413	*Subset van het Handelsregister op basis van een selectie van zorg gerelateerde SBI-codes*
BIG	Registraties (Zorgverleners)	365.385	

Hieronder is de beoogde samenhang schematisch weergegeven.



In hoofdstuk 6 wordt nader ingegaan op de feitelijke samenhang die bij matching van de registers is geconstateerd en op de bijzonderheden die hierbij zijn gebleken.

Onderstaand schema geeft weer op welke wijze de relaties tussen de verschillende registraties is voorzien. Opgemerkt moet worden dat JMV en ZIN in de analyse niet zijn meegenomen. Dit omdat het HR, AGB en BIG de feitelijke basisinformatie leveren over zorgondernemingen en zorgverleners. Het JMV en ZIN sluiten hier vervolgens met andersoortige informatie op aan.



6 Analyse van de samenhang

Bij de analyse van de samenhang van de datasets is nagegaan in hoeverre de beoogde samenhang te realiseren is. Voor de vulling van het LRZa zijn voor de organisaties de gegevens uit Handelsregister de basis, terwijl voor de natuurlijke personen het BIG-register als basis geldt. Voor de gegevens in AGB moet derhalve een overeenkomstige registratie in één van deze bronnen gevonden worden.
De gegevens uit het AGB moeten tegen het Handelsregister en het BIG-register gematcht worden:

- 80.428 zorgverleners zouden moeten voorkomen in het BIG-register;
- 107.949 zorgondernemingen uit AGB zouden moeten voorkomen in het extract vanuit het Handelsregister;
- 115.725 zorgvestigingen uit AGB zouden moeten voorkomen in het extract vanuit het Handelsregister.

Bij de matching is voor ieder van de in het AGB-register opgenomen groepen onderzocht in hoeverre de beoogde matches te realiseren zijn met het Handelsregister en het BIG-register. AGB was dus steeds het uitgangspunt voor de matching.

6.1 Introductie matchingmethodiek

Voor de matching is gebruik gemaakt van software van Human Inference. Deze software maakt gebruik van fouttolerante technieken en kennisdatabases om een optimale interpretatie en beoordeling te realiseren. (zie ook bijlage 5 "Over Human Inference").

De gebruikte software, HIquality Identify, hanteert fouttolerante vergelijkingsmethoden waardoor deze software een oordeel kan geven over de gelijkheid van gegevens zoals een mens dat ook kan doen. Over-kill (teveel treffers) en under-kill (het missen van relevante treffers) worden daarmee tot een minimum teruggebracht. Hoe de vergelijkingen plaatsvinden is vastgelegd in een klantspecifieke configuratie.

Fonologie is één van de onderdelen die bijdraagt aan die fouttolerantie. Immers, er zijn wel 30 verschillende schrijfwijzen van een naam die klinkt als "Matijsen" en die moeten in principe wel bij elkaar gevonden kunnen worden. Anderzijds zijn er ook namen die op meerdere manieren uitgesproken kunnen worden, zoals de naam ""Houben" (uit te spreken als "houben" maar ook als "hoeben").

Vergelijkingsmethoden

Voor ieder gegeven dat vergeleken moet worden is een specifieke vergelijkingsmethode van toepassing. Voor een enkel woord is de methode "woordbeeld" een goede keuze, maar voor een complete tenaamstelling van een organisatie past de methode "matrix-vergelijking" beter. Zo zijn er ook specifieke methodes voor onder meer geboortedatum, huisnummer en postcode.

Wegingen en score

Op basis van onder meer belangrijkheid en vullingsgraad wordt aan ieder gegeven dat vergeleken wordt een gewicht toegekend. Daarmee wordt bepaald in welke mate dat gegeven de overall score (waarmee de overeenkomst wordt aangeduid) beïnvloedt.
Stel dat de naamgegevens 100% overeenkomen en de adresgegevens 20%, dan zou je tot een score van 60% kunnen komen. Door echter de naamgegevens twee keer zwaarder te wegen dan de adresgegevens is de overall score 73%.

Score en beslissingen

In het vergelijkingsproces worden meerdere stappen doorlopen met eigen scorebepaling op (een deel van) de gegevens en bijbehorende beslissingen. Zo kan bijvoorbeeld een beslissing worden gedefinieerd op basis van het handelsregisternummer:

- indien gelijk, altijd een treffer;
- indien afwijkend, nooit een treffer;
- indien onbekend, doorgaan met de volgende stap.

Een volgende stap zou dan bijvoorbeeld een vergelijking van alleen de naamgegevens kunnen zijn etc.

Overall score

Van ieder treffer wordt tot slot een overall score bepaald. Op die manier is eenvoudig te zien of de feitelijke gegevens voldoende overeenkomen bij een treffer die op alleen het handelsregisternummer al gevonden werd. Mogelijke mis-matches op dat nummer komen daarmee snel aan het licht.

In de volgende paragrafen worden per onderdeel de resultaten van de matching toegelicht.

6.2 Zorgondernemingen

Bij de matching van zorgondernemingen zijn de volgende gegevens gehanteerd:

Gegevens	**Weging in de matching**
Naamgegevens	65%
Adresgegevens	35%
Handelsregisternummer	Afzonderlijke beslisregel

Bij het bepalen van de wegingen voor de matching is uitgegaan van een "standaard" configuratie die op basis van de ervaringen van Human Inference als passend beschouwd kan worden. In dit voorbeeld wordt 65% van de overeenkomst bepaald door de tenaamstelling van de organisatie en wordt 35% bepaald door het adres. Dit is een gangbare verhouding wanneer verdere gegevens niet beschikbaar zijn, waarbij het adres bewust een lager gewicht krijgt vanwege verhuizingen of vestiging in een bedrijfs-verzamelgebouw.

De volgende beslisregels zijn gehanteerd:

- Als handelsregisternummer overeenkomt dan is er een match; de score wordt vervolgens bepaald op basis van het totale record;
- Als NAW gegevens een score boven de 75 hebben is er een match.

Hierbij is er rekening mee gehouden dat bepaalde records meerdere adressen kunnen hebben ten opzichte van één adres in het andere record.

De gevonden matches zijn vervolgens onderverdeeld in zekere matches en mogelijke matches. Een zekere match heeft een overall score van 80 en hoger; een mogelijke match heeft een overall score van 75 en hoger, maar lager dan 80.

6.2.1 Benodigde matches

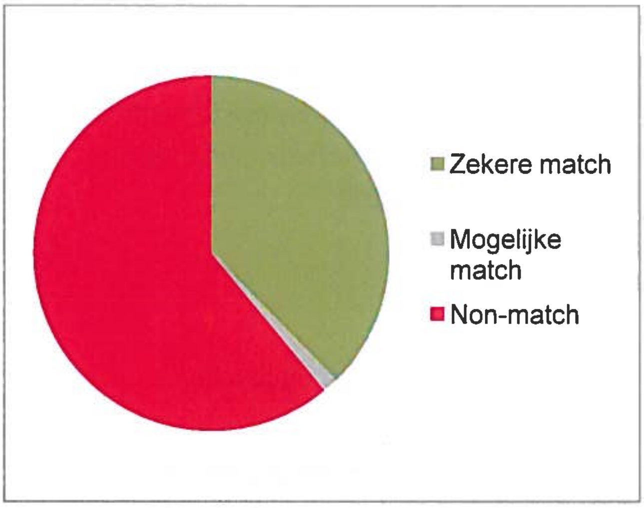
Het totaal aantal zorgondernemingen in AGB bedraagt 107.949, die bekend zouden moeten zijn in het Handelsregister.

6.2.2 Match resultaten

Van de 107.949 zorgondernemingen vanuit AGB is er voor 41.862 een match met het Handelsregister gevonden op basis van de beschreven matchingscriteria; een score van 38,8%.

Verdeling:

- Zekere matches 37,5%
- Mogelijke matches 1,3%
- Non-matches 61,2%



Bij een matching op alleen het handelsregistratienummer bedraagt het aantal matches 3.804, ofwel 3,5%. Verrijking van AGB met het handelsregisternummer zal prioriteit moeten hebben om de matching te verbeteren en de kwaliteit van het LRZa te verhogen.

Er zijn echter ook AGB-records waarvoor het handelsregisternummer is ingevuld, maar toch geen match met het Handelsregister is gevonden omdat ze niet voorkomen in de aangeleverde subset. In de tabel hieronder zijn van enkele gevallen de resultaten van een nadere analyse opgenomen:

Handelsregnr	Tenaamstelling	Opmerkingen
09221030	VERPLEEGHUIS HEIJENDAAL	Tenaamstelling volgens het Handelsregister is "Stichting DrieGasthuizenGroep". VERPLEEGHUIS HEIJENDAAL is ook bekend in het Handelsregister maar onder nummer 41051592.

Handelsregnr	Tenaamstelling	Opmerkingen
34152255	INTERAPY NEDERLAND B.V. (REGIO AMSTERDAM)	Het aangeleverde handelsregisternummer is uitgeschreven. Het juiste handelsregister-nummer is 34172303.
24298945	WAARDEBURGH PARKZICHT (VZH)	Het aangeleverde handelsregisternummer is niet bekend. Het juiste handelsregister-nummer is 24299845.
41064097	STICHTING THUISZORG NOORD-LIMBURG	Het aangeleverde handelsregisternummer is uitgeschreven.
27250988	STICHTING ZORGGROEP FLORENCE (THUISZORG)	Het aangeleverde handelsregisternummer is correct, maar de betreffende vermelding is niet opgenomen in de aangeleverde subset.
14090292	HUIZE DE BERG	Het aangeleverde handelsregisternummer is uitgeschreven. Het juiste handelsregister-nummer is 14090032.

6.2.3 Opvallende resultaten

Bij de analyse van de matchingresultaten zijn een aantal resultaten opgevallen.

Bij 3.804 records uit AGB is er sprake van een 100% match op handelsregisternummer, maar is overeenkomst van de tenaamstelling en adresgegevens aanmerkelijk kleiner. Bij nadere beschouwing lijkt de conclusie terecht dat het in veel van deze gevallen gaat om vestigingen in plaats van ondernemingen.
Een sprekend voorbeeld is "Zorgpartners Friesland" (zie hieronder), waarbij de hoofdvestiging die bij het Handelsregister is geregistreerd in AGB niet voorkomt, maar 21 vestigingen wel als zorgonderneming zijn geregistreerd terwijl dat zorgvestigingen zouden moeten zijn. Die vestigingen komen echter in het Handelsregister weer niet voor als vestigingen van "Zorgpartners Friesland".

Bron	Organisatienaam	KvK-nr	Straat	Huisnr	Postcode	Plaats
HR	Zorgpartners Friesland	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (ERASMUSHIEM VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (MECKAMA STATE VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]

Bron	Organisatienaam	KvK-nr	Straat	Huisnr	Postcode	Plaats
AGB	STICHTING ZORGPARTNERS FRIESLAND (NIEUW MELLENS VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (DE STILEN VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (DE SPIKER VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (NIEUW TOUTENBURG VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (BENNEMA STATE VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (DE STILEN VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (MECKAMA STATE VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (ABBINGAHIEM VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (NIJ BETHANIE VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (DE BATTING VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (BORNIA HERNE VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (DE SPIKER VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (FRIESMA STATE VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (THUISZORG NOORDERBREEDTE)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (ERASMUSHIEM VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (BENNEMA STATE VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (NIJ BETHANIE VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (ABBINGAHIEM VZH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]
AGB	STICHTING ZORGPARTNERS FRIESLAND (FRIESMA STATE VPH)	1079054	[redacted]	[redacted]	[redacted]	[redacted]

Een ander voorbeeld betreft "A S Watson" wat de moedermaatschappij is van onder meer "Kruidvat" en "Trekpleister". Allen (1.020 registraties in het Handelsregister) hebben hetzelfde handelsregisternummer, maar zijn afzonderlijke organisaties en/of vestigingen.

Bij het gebruik van het handelsregisternummer om de koppeling tussen AGB en Handelsregister te leggen is een toetsing van de werkelijke overeenkomst van de gegevens onontbeerlijk. Verder is opname van het vestigingsnummer in AGB noodzakelijk om de juiste registratie in het Handelsregister te kunnen bepalen en de kwaliteit van de matching te verhogen.

Daarnaast is bij een tweetal zorgsoorten onderzocht hoeveel van de in het AGB-register opgenomen zorgondernemingen een match hebben met het Handelsregister.

Zorgsoort	**Aantal AGB-registratie**	**Percentage matches met Handelsregister**
Tandartsen	6768	45,5%
Huisartsen	5644	51,2%

Een conclusie zou kunnen zijn dat een groot deel van deze AGB-registraties niet is ingeschreven in het Handelsregister danwel dat ze niet in de aangeleverde subset zijn opgenomen.
Een aantal steekproeven voor deze zorgsoorten hebben aangegeven dat het veelal zal gaan om tandartsen en huisartsen die in AGB-register als zorgonderneming zijn geregistreerd, maar feitelijk in een maatschap samenwerken. Die maatschap is wel ingeschreven in het Handelsregister, maar onbekend in het AGB-register.
Een logische conclusie lijkt dat:

1. de maatschap alsnog in het AGB-register van zorgondernemingen opgenomen moet worden;
2. de betreffende huisartsen en tandartsen niet als zorgonderneming, maar als zorgvestiging c.q. zorgverlener geregistreerd moeten worden;
3. het Handelsregister, in het geval van een registratie als zorgvestigingen, deze vestigingen ook onder de geregistreerde maatschap zal moeten opnemen.

Hieronder is een voorbeeld weergegeven van hoe huisartsen in Duiven geregistreerd zijn in het Handelsregister en in AGB. Deze huisartsen werken alle in dezelfde maatschap samen. Die maatschap is als enige geregistreerd in het Handelsregister, maar de vestigingen en afzonderlijke huisartsen ontbreken daar. In AGB ontbreekt echter de maatschap en zijn alleen de afzondelijke praktijken geregistreerd.
Deze wijze van registreren kan gevolgen hebben voor een eenduidige registratie in het LRZa.

Handelsregister

Handels-registernr	Organisatienaam	Straat / Huisnummer	Postcode / Plaats
09220445	Huisartsenmaatschap Duiven	[redacted]	[redacted]

AGB

Zorgpartijid	Organisatienaam	Straat / Huisnummer	Postcode / Plaats
1242929	HUISARTSENPRAKTIJK FLOWER	[redacted]	[redacted]
1203448	JANSEN HUISARTSENPRAKTIJK	[redacted]	[redacted]
1203447	REICHERT/REICHERT-RONDA/DELLEN-COSTER	[redacted]	[redacted]
1238264	HUISARTSENPRAKTIJK JONGEN & VOORHAAR	[redacted]	[redacted]
1203434	SLUIJMERS/LIEM/ROMMENS HUISARTSENPRAKTIJK	[redacted]	[redacted]
1216855	MULLING	[redacted]	[redacted]

6.3 Zorgvestigingen

Bij de matching van zorgvestigingen zijn de volgende gegevens gehanteerd:

Gegevens	Weging in de matching
Naamgegevens	65%
Adresgegevens	35%
Handelsregisternummer	Afzonderlijke beslisregel

Het vestigingsnummer kon vanwege de lage vullingsgraad niet gebruikt worden voor de matching.

De volgende beslisregels zijn gehanteerd:

- Als handelsregisternummer overeenkomt dan is er een match; de score wordt vervolgens bepaald op basis van het totale record;
- Als NAW gegevens een score boven de 75 hebben is er een match.

Hierbij is er rekening mee gehouden dat bepaalde records meerdere adressen kunnen hebben ten opzichte van één adres in het andere record.

De gevonden matches zijn vervolgens onderverdeeld in zekere matches en mogelijke matches. Een zekere match heeft een overall score van 80 en hoger; een mogelijke match heeft een overall score van 75 en hoger, maar lager dan 80.

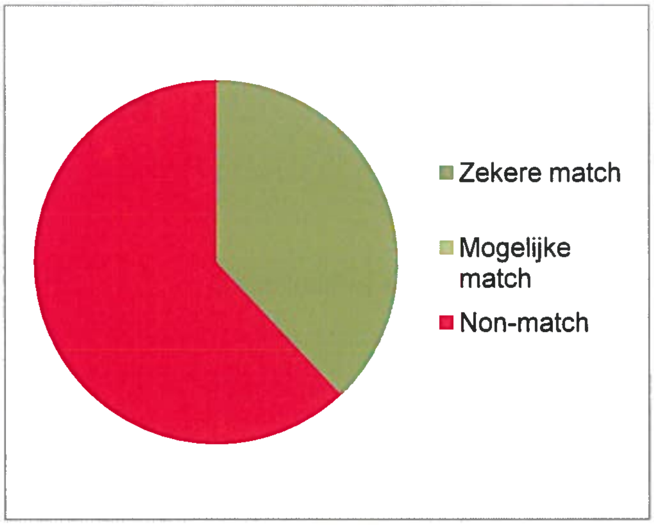
6.3.1 Benodigde matches

Het totaal aantal zorgvestigingen in AGB bedraagt 115.725, die bekend zouden moeten zijn in het Handelsregister.

6.3.2 Match resultaten

Van de 115.725 zorgvestigingen vanuit AGB is er voor 43.693 een match met het Handelsregister gevonden op basis van de beschreven matchingscriteria; een score van 37,8%. Verdeling:

- Zekere matches 37,8%
- Mogelijke matches 0,0%
- Non-matches 62,2%



Bij een matching op alleen het vestigingsnummer bedraagt het aantal matches 44: een bijna te verwaarlozen aantal. De vulling van het vestigingsnummer in AGB, maar ook de registratie van vestigingen in het handelsregister, is onvoldoende om basis van het vestigingsnummer te matchen.

6.3.3 Opvallende resultaten

Bij de analyse van de matchingresultaten zijn een aantal resultaten opgevallen.
Bij 44 records uit AGB is er sprake van een 100% match op vestigingsnummer, maar is in veel gevallen de overeenkomst van de feitelijke gegevens lager dan 70%. Dit wordt met name veroorzaakt door afwijkende tenaamstellingen, zoals hieronder te zien is:

Vestigingsnr	Tenaamstelling Handelsregister	Tenaamstelling AGB
000028015754	CBS Klingelenburg	LOGOPEDIEPRAKTIJK BEEKWOUT
000023107391	OBS Meester Aafjesschool	LOGOPEDIEPRAKTIJK BEEKWOUT
000006175198	Estafette	ACUPUNCTUUR HET GOOI
000023237163	Fysio4All B.V.	FYSIOTHERAPIE CINDY BOGERT
000020058144	Fysiotherapie 'De La Reyweg 259'	STICHTING FYSIOCONSULT-DEN HAAG
000027739120	ISV (Europe) B.V.	APOTHEEK VITA NATURA
000028015754	Raadhuis Krimpen aan den IJssel	MUZIEKTHERAPIE ANTOINETTE MUZIEKSCHOOL KRIMPEN AAN DEN IJSSE

Logopediepraktijk Beekwout blijkt in AGB meerdere keren opgenomen te zijn met telkens een ander vestigingsnummer. Die vestigingsnummers zijn echter het vestigingsnummer van de betreffende school waarin de praktijk mogelijk zitting houdt en niet een vestigingsnummer dat bij de logopediepraktijk hoort.

6.4 Zorgverleners

Bij de matching van zorgverleners zijn de volgende gegevens gehanteerd:

Gegevens	Weging in de matching
Naamgegevens	32%
Adresgegevens	39%
Geboortedatum	14%
Geslacht	7%
Telefoonnummer	4%
E-mail adres	4%

Bij het bepalen van de wegingen voor de matching is uitgegaan van een "standaard" configuratie die op basis van de ervaringen van Human Inference als passend beschouwd kan worden.

In dit voorbeeld wordt 53% van de overeenkomst bepaald door persoonsgebonden gegevens (naam, geboortedatum, geslacht), wordt 39% bepaald door het adres (iemand kan verhuizen) en wordt 8% bepaald door persoonsgerelateerde gegevens (telefoonnummer, e-mail adres).

De volgende beslisregels zijn gehanteerd:

- Als achternaam, geboortedatum, geslacht en voorletters gelijk zijn (score 100) dan is er een match;
- Als NAW gegevens een score boven de 85 hebben én geboortedatum overeenkomt of bij één van de records leeg is dan is er een match;
- Als de gegevens een overall score van 75 en hoger hebben dan is er een match.

Hierbij is er rekening mee gehouden dat bepaalde records meerdere adressen kunnen hebben ten opzichte van één adres in het andere record.

De gevonden matches zijn vervolgens onderverdeeld in zekere matches en mogelijke matches. Een zekere match heeft een overall score van 80 en hoger; een mogelijke match heeft een overall score van 75 en hoger, maar lager dan 80.

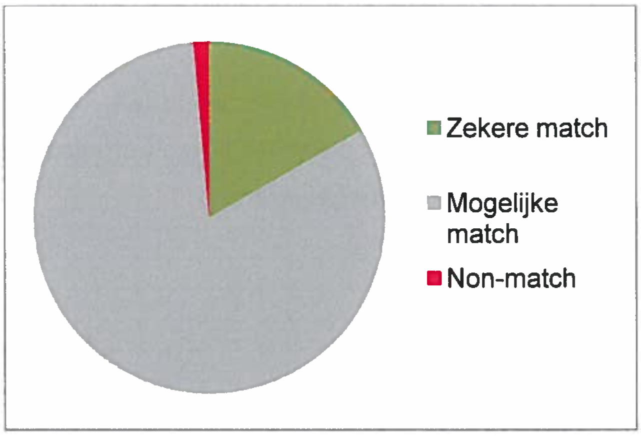
6.4.1 Benodigde matches

Het totaal aantal zorgverleners in AGB bedraagt 163.065.
Vanuit AGB gezien zouden er 80.428 zorgverleners bekend moeten zijn in het BIG-register aangezien deze voldoen aan de zorgsoorten die een BIG registratie moeten hebben en waarbij er geen einddatum bekend is voor de AGB-code.

6.4.2 Match resultaten

Van de 80.428 zorgverleners vanuit AGB is er voor 79.303 een match met het BIG-register gevonden op basis van de beschreven matchingscriteria; een score van 98,6%.
Verdeling:

- Zekere matches 16,7%
- Mogelijke matches 81,9%
- Non-matches 1,4%



Aangezien bij de door Human Inference uitgevoerde matching de score op basis van alle kenmerken wordt bepaald, valt hieruit te concluderen dat de NAW-gegevens van de registraties in het BIG-register en in AGB vaak niet overeenkomstig zijn, hetgeen voor een deel veroorzaakt wordt door het ontbreken van adresgegevens (werkadres) bij ruim 41% van de registraties in het BIG-register.
Op basis van het BIG-registratienummer wordt slechts 16,7% als voldoende gelijk en daarmee als zekere match gekwalificeerd. Synchronisatie en aanvulling van de NAW-gegevens lijkt noodzakelijk, bijvoorbeeld door het aanvullen van het werkadres in BIG vanuit AGB.

Een tweetal voorbeelden is hieronder opgenomen:

Bron	**Naam**	**Geboortedatum**	**Geslacht**
BIG-regnr	**Adres**	**Telefoonnr**	**E-mail**
AGB	MT VAN REISEN	[redacted]	[redacted]
19053709001	[redacted]	[redacted]	[redacted]
BIG	M.T. Ruiter-van Reisen	...	...
19053709001	...	...	...
AGB	MR FABER	[redacted]	[redacted]
19053718901	[redacted]	[redacted]	[redacted]
BIG	M.R. de Bruin-Faber	...	...
19053718901	...	...	...

7 Verrijkingsmogelijkheden

In de voorgaande hoofdstukken (4 en 6) zijn opmerkingen gemaakt over vullingsgraad, over volledigheid, over kwaliteit van de gegevens en over betrouwbaarheid van gehanteerde koppelingen op basis van nummers.
Door de thans bekende gegevens te verrijken ontstaat een hogere kwaliteit, kunnen de gegevens beter gematcht worden en ontstaat daarmee een kwalitatief hoogwaardig en betrouwbaar LRZa.

Verrijkingsmogelijkheden liggen deels in de reeds bekende data opgesloten, maar behoeven veelal ook externe bronnen waarin aanvullende informatie bekend is. Primair zullen die verrijkingen in de bronsystemen moeten plaatsvinden, bij voorkeur voorafgaand aan de vulling van het LRZa. Met behulp van de HIquality software van Human Inference kan verrijking onderdeel zijn van de vulling van het LRZa.

7.1 AGB Zorgonderneming

Het handelsregisternummer is op dit moment zeer beperkt gevuld, zodat een verdere vulling hiervan prioriteit zou moeten hebben. Immers, deze nummers zijn voor de vulling van het LRZa belangrijk.
Met behulp van HIquality Identify kan de data van AGB effectief gematcht worden tegen het volledige Handelsregister, waarbij de unieke treffer (met voldoende overeenkomst) gekozen wordt dan wel de mogelijke treffers gepresenteerd worden. Hiermee kan een correcte vulling van AGB relatief snel worden gerealiseerd. Het risico van verkeerde koppelingen is met de juiste instellingen uit te sluiten, zonder dat een 100% gelijkheid vereist is.

7.2 AGB Zorgvestigingen

Het handelsregisternummer en het vestigingsnummer zijn op dit moment zeer beperkt gevuld, zodat een verdere vulling hiervan prioriteit zou moeten hebben. Immers, deze nummers zijn voor de vulling van het LRZa belangrijk.
Met behulp van HIquality Identify kan de data van AGB effectief gematcht worden tegen het volledige Handelsregister, waarbij de unieke treffer (met voldoende overeenkomst) gekozen wordt dan wel de mogelijke treffers gepresenteerd worden. Hiermee kan een correcte vulling van AGB relatief snel worden gerealiseerd. Het risico van verkeerde koppelingen is met de juiste instellingen uit te sluiten, zonder dat een 100% gelijkheid vereist is.

7.3 AGB Zorgverleners

Bij 1,3% van de registraties ontbreken de voorletters als afzonderlijke rubriek, maar zijn die voorletters wel opgenomen als onderdeel van de naam.
Met behulp van HIquality Name kunnen deze voorletters eenvoudig worden bepaald en in de juiste rubriek worden opgenomen.

Het BIG-registratienummer is op dit moment gevuld bij 99,5% van de zorgverleners die een BIG-registratie zouden moeten hebben. Met behulp van HIquality Identify kan een aanvullende matching tussen AGB en BIG worden uitgevoerd om ook voor de ontbrekende registraties het BIG-registratienummer te kunnen herleiden. Het risico van verkeerde koppelingen is met de juiste instellingen uit te sluiten, zonder dat een 100% gelijkheid vereist is.

7.4 Handelsregister

We hebben de indruk dat in AGB vestigingen opgenomen die in het Handelsregister op dit moment nog niet bekend zijn. Nadat de matching van AGB met het Handelsregister is voltooid kan die informatie aan het Handelsregister worden geleverd. Op basis van eigen onderzoek zal het Handelsregister vervolgens kunnen bepalen of opname van die vestigingen gewenst is en de daarbij behorende acties kunnen ondernemen.

Op basis van de thans gevonden matches tussen AGB en het Handelsregister bestaat de indruk dat voor de initiële vulling van het LRZa de subset op basis van SBI-codes geen volledige dekking biedt. Het is daarom aan te bevelen om ten behoeve van de initiële vulling van het LRZa een volledige extractie uit het Handelsregister als referentieset beschikbaar te hebben.

7.5 BIG-register

Bij ruim 41% van de registraties is geen werkadres bekend. Op basis van de GBA zou het woonadres meegeleverd kunnen worden, maar dat is voor de vulling van het LRZa op dit moment niet toegestaan. Via de GBA zouden ook de ontbrekende voorletters kunnen worden aangevuld, echter ook dat is op dit moment niet toegestaan.

7.6 Prioritering

In de opbouw van het LRZa spelen de gegevens uit AGB een centrale rol. Die gegevens van zorgondernemingen en zorgvestigingen moeten immers gematcht worden met het Handelsregister. En voor de zorgverleners geldt dat een groot deel gematcht moet worden met het BIG-register.
Het is daarom aanbevolen de verrijkingsmogelijkheden m.b.t. AGB prioriteit te geven.

8 Conclusies en aanbevelingen

8.1 Conclusies

In de voorgaande hoofdstukken zijn de resultaten van de uitgevoerde analyses weergegeven, waarbij bevindingen en opmerkelijke resultaten zijn benoemd. Op basis van die analyses zijn in deze paragraaf de belangrijkste conclusies samengebracht.

8.1.1 Aangeleverde datasets

- Data wordt vanuit de verschillende bronnen in verschillende formaten aangeleverd, wat bij exacte vergelijking eenvoudig tot non-matches kan leiden;
- De verschillende formaten van aanlevering, waaronder één of meerdere velden voor naamgegevens, zijn een nadeel voor een correcte vulling van het LRZa;
- De vullingsgraad van AGB is op een aantal onderdelen onvoldoende:
  - Bij slechts 3,9% van de zorgondernemingen (4.262 van 107.949) is een handelsregisternummer geregistreerd, terwijl dit nummer een belangrijke rol zou moeten spelen bij de vulling van het LRZa;
  - Geregistreerde handelsregisternummers zijn in een aantal gevallen niet correct;
  - Bij slechts 61 zorgvestigingen is een handelsregistervestigingsnummer geregistreerd, terwijl dit nummer een belangrijke rol zou moeten spelen bij de vulling van het LRZa;
  - Bij 1,3% van de zorgverleners zijn geen voorletters geregistreerd, maar zijn die mogelijk wel af te leiden uit de tenaamstelling;
  - Bij 0.5% van de zorgverleners is geen BIG-registratienummer geregistreerd;
- In het BIG-register is bij 41% van de registraties geen werkadres bekend;
- De vanuit het Handelsregister aangeleverde subset van inschrijvingen op basis van geselecteerde SBI-codes lijkt te beperkt, want zelfs met fouttolerante matching kan voor slechts 38,8% van de zorgondernemingen een match gevonden worden.

8.1.2 Datakwaliteit

Over het geheel genomen kan worden gesteld dat de kwaliteit van de tenaamstellingen, adresseringen, telefoonnummers en e-mail adressen goed is. Een gemiddeld percentage variërend van 1% tot 5% aan fouten en te beoordelen constateringen vereist echter wel aandacht en maatregelen om te komen tot een kwalitatief hoogwaardig en betrouwbaar LRZa.

8.1.3 Matching AGB-Handelsregister

Zorgondernemingen:

- Voor slechts 38,8% van de zorgondernemingen in AGB kon een match met het Handelsregister gevonden worden;
- Een aantal in AGB geregistreerde handelsregisternummers is niet correct;
- Exacte matching van gegevens uit de verschillende bronnen zal leiden tot veel non-matches als gevolg van afwijkende schrijfwijzen;
- Een aantal registraties in AGB waarbij een handelsregisternummer bekend is lijkt feitelijk als zorgvestiging geregistreerd te moeten worden in plaats van als zorgonderneming;
- De aanwezigheid in het Handelsregister van alle zorgondernemingen die in het AGB zijn opgenomen lijkt, gelet op de analyse die als steekproef voor de zorgsoorten "Tandartsen" en "Huisartsen" is gemaakt, niet gegarandeerd te kunnen worden.

Zorgvestigingen:

- Voor slechts 37,8% van de zorgvestigingen in AGB kon een match met het Handelsregister gevonden worden;
- Exacte matching van gegevens uit de verschillende bronnen zal leiden tot veel non-matches als gevolg van afwijkende schrijfwijzen (b.v. "TROMBOSEDIENST VAN HET ARTSENLABORATORIUM VOOR DE ZAANSTREEK" versus "Stg. Artsenlaboratorium en Trombosedienst");
- Een aantal in AGB geregistreerde handelsregistervestigingsnummers is niet correct.

8.1.4 Matching AGB-BIG

- Voor 98,6% van de zorgverleners in AGB kon een match met het BIG-register gevonden worden;
- Bij die AGB-registraties waar het BIG-registratienummer gevuld is, is in 81,4% van de gevallen sprake van een beperkte overeenkomst op de feitelijke gegevens (door het ontbreken van adresinformatie (werkadres) in BIG), wat de betrouwbaarheid van de matches nadelig beïnvloedt;
- Exacte matching van gegevens uit de verschillende bronnen zal leiden tot veel non-matches als gevolg van afwijkende schrijfwijzen (b.v. "MT VAN REISEN" versus "M.T. Ruiter-van Reisen").

8.2 Aanbevelingen

Op grond van de uitgevoerde analyse en ervaringen van Human Inference in soortgelijke trajecten komen wij tot een aantal aanbevelingen die naar onze mening zullen leiden tot een kwalitatief hoogwaardig en betrouwbaar LRZa.

8.2.1 Optimalisatie databronnen

AGB-register

- Zorg voor een volledige en correcte vulling van het handelsregisternummer bij de zorgondernemingen op basis van een intelligente matchingmethode, waardoor de reeds door Vektis gedefinieerde actie wordt versneld;
- Zorg voor een volledige en correcte vulling van het handelsregisternummer en vestigingsnummer bij de zorgvestigingen op basis van een intelligente matchingmethode, waardoor de reeds door Vektis gedefinieerde actie wordt versneld;
- Analyseer de meldingen met betrekking tot de datakwaliteit en neem maatregelen ter verbetering van het AGB-register (ook wanneer de zorgaanbieders zelf verantwoordelijk zijn voor de registratie van hun gegevens).

Handelsregister

- Zorg voor een verdere vulling van het Handelsregister met vestigingen, bijvoorbeeld op basis van informatie vanuit het AGB-register.
  Op basis van eigen onderzoek zal het Handelsregister vervolgens kunnen bepalen of opname van die vestigingen gewenst is en de daarbij behorende acties kunnen ondernemen.

8.2.2 Aanlevering van data

- Hanteer voor de initiële vulling van het LRZa de volledige dataset van het Handelsregister als referentie in plaats van een subset op basis van SBI-codes; in LRZa worden uiteindelijk alleen de relevante vermeldingen uit het Handelsregister opgenomen.

8.2.3 Verwerking in LRZa

- Richt een proces en structuur in waarbij de vanuit de databronnen aangeleverde gegevens worden gestandaardiseerd, zodat de gegevens (ongeacht de bron) eenduidig kunnen worden gematcht. De gegevens uit de authentieke bronnen blijven in LRZa leidend.

8.2.4 Matching

- Hanteer bij de matching op exacte criteria (Handelsregisternummer, BIG-registratienummer) ook fouttolerante matchingsoftware voor de feitelijke data om de mate van overeenkomst vast te stellen en onjuiste koppelingen tijdig te signaleren (zgn. "naam-nummer-check");
- Hanteer fouttolerante matchingsoftware voor matches die niet op basis van exacte criteria gevonden kunnen worden, zoals bij meerdere registraties met hetzelfde handelsregisternummer.
- Hanteer intelligente zoeksoftware (dezelfde HIquality Identify software) binnen de websites die gebruikt worden om het LRZa te raadplegen, zodat de gezochte vermeldingen eenvoudig gevonden worden; hiervoor wordt één uniforme webservice toegepast.

8.2.5 Vulling LRZa

- Uitgangspunt voor de vulling van het LRZa zijn de gegevens van het Handelsregister en het BIG-register, maar verrijking op basis van gegevens uit AGB is alleen mogelijk mits goede samenvoegregels worden gehanteerd;
- Koppel zorgondernemingen die vanuit het AGB-register worden aangeleverd en waarbij de overeenkomst van de feitelijke gegevens met het Handelsregister te laag is niet aan de organisatie, maar registreer die in het LRZa als vestiging onder die organisatie;
- Koppel vestigingen die vanuit het AGB-register worden aangeleverd en waarbij de overeenkomst van de feitelijke gegevens met het Handelsregister te laag is niet aan die vestiging, maar registreer die in het LRZa als nieuwe vestiging;
- Signaleer onvoldoende overeenkomende gegevens ten behoeve van terugmelding aan de bronhouder.

8.2.6 Operationeel gebruik LRZa

Tot slot willen wij nog enkele aanbevelingen geven met betrekking tot het operationeel gebruik van het LRZa.

- Bij het gebruik van het LRZa is het essentieel dat de opgenomen informatie toegankelijk is en geregistreerde zorgondernemingen, zorgvestigingen en zorgverleners eenvoudig gevonden kunnen worden. Een fouttolerante zoekmethode zorgt ervoor dat de juiste registraties gevonden worden, maar kan zo worden geconfigureerd dat het aantal resultaten beperkt blijft;
- Hanteer het LRZa als een toepassing om gegevens te ontsluiten, maar hanteer voor de verwerking van gegevens vanuit de bronsystemen een datakwaliteits-hub voor validatie, standaardisatie en matching;
- Zorg voor monitoring functies als onderdeel van die datakwaliteits-hub zodat een goed inzicht bestaat in de kwaliteit van de aangeleverde informatie.

Datakwaliteits-hub
Een datakwaliteits-hub is een bundeling van datakwaliteitsfuncties (services) voor het valideren en standaardiseren van gegevens en zorgt daarmee voor borging en verhoging van de kwaliteit. Verder zijn services opgenomen voor de matching en het fouttolerant zoeken om enerzijds de juiste gegevens eenvoudig te vinden en anderzijds de juiste koppelingen te leggen.

8.2.7 Prioritering

Prioriteit 1
In de opbouw van het LRZa spelen de zorgondernemingen en zorgvestigingen een centrale rol. Prioriteit voor de optimalisatie van de gegevens uit AGB ten behoeve van de matching met het Handelsregister is daarom aanbevolen.
In eerste aanleg is het aanbevolen om alle in het AGB geregistreerde ondernemingen aan te schrijven om de gegevens te completeren danwel te actualiseren. In tweede aanleg zou dan een geautomatiseerde matching van de AGB registratie met het Handelsregister uitgevoerd kunnen worden, zodat daarmee de gegevens in AGB worden geoptimaliseerd. Om deze werkwijze in de praktijk te toetsen is een pilot voor een bepaalde regio (bijvoorbeeld een bepaald postcodegebied) aanbevolen.

Prioriteit 2

Tussen de aanbevelingen zoals genoemd in paragraaf 8.2.3 t/m 8.2.6 bestaat een grote samenhang ten aanzien van de initiële opbouw van het LRZa en het gebruik van het LRZa. Ook om deze aanbevelingen in de praktijk te toetsen is een pilot voor een bepaalde regio (bijvoorbeeld een bepaald postcodegebied) aanbevolen. Wanneer voor dezelfde regio wordt gekozen als bij prioriteit 1, wordt tevens een gefaseerde vulling van het LRZa gefaciliteerd.

Prioriteit 3

Op basis van de resultaten van de pilot kan de definitieve opzet, de (verdere) vulling en het gebruik van het LRZa nader worden gedefinieerd, zodat de realisatie van het volledige LRZa voorspoedig zal verlopen.

8.3 Advies high-level dataflow

Onderstaande dataflow is een weergave van de, naar de mening van Human Inference, beste aanpak voor de vulling van het LRZa. Aangezien de levering van de data in verschillende formaten plaatsvindt en koppeling niet alleen op basis van nummers mogelijk is levert de HIquality software een wezenlijke meerwaarde door te fungeren als een datakwaliteits-hub.



HIquality software van Human Inference is gebaseerd op jarenlange research en ervaring op het gebied van de inhoudelijke kwaliteit van relatiegegevens (naam, adres, telefoonnummer, email adres etc.) alsmede op het gebied van fouttolerante matching van gegevens. Het helpt organisaties om data vanuit verschillende bronnen te valideren, te standaardiseren, te matchen en te consolideren in een centraal raadpleegbaar register.

Praktijkervaringen uit implementaties, binnen de overheid bij onder meer de Belastingdienst, Dienst Uitvoering Onderwijs, Justitie, Politie, Kadaster en Kamer van Koophandel, tonen het belang van fouttolerante matching en van validatie en standaardisatie van gegevens aan.

De hierboven getoonde dataflow schetst de levering van gegevens vanuit de bronnen aan de datakwaliteits-hub ten behoeve van de vulling van het LRZa. In relatie tot het LRZa worden door gebruik te maken van de HIquality software:

- de gegevens uit de verschillende bronnen eerst gematcht op basis van unieke nummers (handelsregisternummer, vestigingsnummer, BIG-registratienummer);
- de gegevens uit de verschillende bronnen vervolgens gematcht op basis van de feitelijk aangeleverde gegevens zoals de NAW-gegevens;
- de gegevens uit de verschillende bronnen gecontroleerd op datakwaliteit t.a.v. de NAW-gegevens;
- de gegevens uit de verschillende bronnen gestandaardiseerd, zodat het LRZa de gegevens eenduidig kan kan weergeven aan een raadpleger. Belangrijk hierbij te noemen is dat de software de gegevens in de bronnen niet aanpast. Dit betekent dus dat het LRZa een eenduidige weergave van records kan hanteren terwijl de bronnen hun eigen notatie handhaven;
- indien gewenst, de aangeboden en gematchte records samengevoegd tot een "golden" record op basis van verschillende regels:
    - Bron prioriteit;
    - Laatst gewijzigd;
    - Hoogste of laagste nummer;
    - Kwaliteit van de gegevens;
    - Etc.;

Deze regels kunnen worden ingesteld per veld of groep van velden, zodat elke mogelijk combinatie mogelijk is. En er kan een keten van regels worden toegepast, zodat er altijd een keuze gemaakt kan worden voor het samenvoegen van de gegevens.

8.3.1 Resumé

Het inzetten van HIquality software zal naar onze mening en ervaring leiden tot een kwalitatief hoogwaardig LRZa. Door naast het matchen op unieke nummers ook (fouttolerant) te toetsen op overige aangeleverde gegevens zal de kwaliteit van het LRZa al bij de vulling toenemen.

Daarnaast vormt dezelfde HIquality software in het beheren van het LRZa een controlemechanisme dat naast de bestaande matches, overige aangeleverde gegevens kan matchen op fouttolerante wijze waardoor meer inzicht wordt verkregen in de samenhang van deze gegevens.

Ook de kwaliteit van de weergave van het LRZa aan de raadpleger zal toenemen doordat de HIquality software de aangeleverde gegevens in één gekozen formaat weergeeft, zonder dat de bronnen hun eigen notatiewijze hoeven aan te passen. Dit zal ook niet leiden tot extra terugmeldingen aan de bronnen.

Tot slot zorgt een fouttolerante zoekmethode ervoor dat door de raadpleger de juiste registraties eenvoudig gevonden worden, zonder dat het aantal resultaten enorm groot wordt.

Bijlage 1 Hoofdactiviteiten AGB

Zorgverleners

Zorgsoort	**Aantal**
Apothekers	2866
Bedrijfsartsen	1918
Diëtisten	3752
Ergotherapeuten	2656
Fysiotherapeuten	28899
Genezers	26148
Huisartsen	11166
Logopedisten	4993
Medisch Specialisten	21328
Mondhygiënisten	1198
Oefentherapeuten	2055
Optometristen	18
Overige Artsen	9194
Pedicuren	11337
Physician Assistent	87
Podotherapeuten	792
Psychologische Zorgverleners	15870
Schoonheidsspecialistes	2787
Tandarts - Specialisten (Dento - Maxillaire Orthopedie)	329
Tandarts - Specialisten (Mondziekten en Kaakchirurgie)	301
Tandartsen	8414
Tandtechnici / Tandprothetici	515
Verloskundigen	2881
Verpleegkundigen	3571

Zorgondernemingen

Zorgsoort	Aantal
Abortusklinieken	45
Ambulancediensten	85
Apothekers	1016
Arbodiensten	169
Arts Arbeid en Gezondheid/Bedrijfsgeneeskunde	527
Audiologische Centra	21
AWBZ Gecombineerd	1495
Beheerst. Gezinsv. Tehuizen, Dagverblijven en Zwakz Inst.	324
Beheerstichtingen Dagverblijven	30
Beheerstichtingen Verpleging,Verzorging en Thuiszorg	406
Beheerstichtingen Verzorgingstehuizen	114
Bloedbanken	22
Consultatiebureaus voor Alcohol en Drugs	5
CPA's	28
Dagverblijf voor Ouderen	199
Declaranten	621
Dialyse-Centra	3
Dienstenstructuren (ANW-Diensten)	60
Dietisten	1533
Diverse Samenwerkingsverbanden	562
Ergotherapeuten	574
Eurotransplant	1
Fysiotherapeuten	8626
Gecombineerde Verpleeginrichtingen	910
Genezers	24791
Gezinsvervangend Tehuis	413
Gezondheidscentra	33
GGD	35
GGZ Instellingen (PUK/PAAZ)	53
Huisartsen	5644
Inrichting voor Auditief Gehandicapten	30
Inrichting voor Visueel Gehandicapten	8
Instellingen voor Jeugdtandverzorging	19
Kinderdagverblijf	101
Klinisch-Genetische Centra	9
Kraamcentra	520

Zorgondernemingen

Zorgsoort	Aantal
Laboratoria(Huisartsenlab./Gemeensch.Lab/Gemeensch Apoth+Lab	74
Leveranciers Hulpmiddelen	8123
Logopedisten	2152
Medisch Specialisten	1451
Mondhygienisten	925
Oefentherapeuten	1753
Optometristen	7
Overige Artsen	2621
Overige Instellingen	139
Pedicuren	11049
Podotherapeuten	437
Psychologische Zorgverleners	8422
R.I.A.G.G. Instelling	52
Radiotherapeutische Centra	6
Rechtspersonen	1286
RIBW	159
Schoonheidsspecialistes	2393
Sociaal Pedagogische Diensten	29
Tandarts-Specialisten (Dento-Maxillaire Orthopedie)	308
Tandarts-Specialisten (Mondziekten en Kaakchirurgie)	37
Tandartsen	6768
Tandheelkundige Centra	536
Tandtechnici/Tandprothetici	566
Taxivervoerders	1486
Thuiszorginstelling	1480
Trombosediensten	44
Verloskundigen	823
Verpleeginrichtingen voor Psycho-Geriatrische Patienten	47
Verpleeginrichtingen voor Somatische Ziekten	36
Verpleegkundigen	2264
Verzorgingshuizen	1516
Zelfstandige Behandelcentra	1101
Ziekenhuizen	514
Zwakzinnigen	311

Zorgvestigingen

Zorgsoort	Aantal
Abortusklinieken	45
Ambulancediensten	85
Apothekers	2201
Arbodiensten	169
Arts Arbeid en Gezondheid/Bedrijfsgeneeskunde	548
Audiologische Centra	21
AWBZ Gecombineerd	1495
Beheerst. Gezinsv. Tehuizen, Dagverblijven en Zwakz Inst.	324
Beheerstichtingen Dagverblijven	30
Beheerstichtingen Verpleging,Verzorging en Thuiszorg	406
Beheerstichtingen Verzorgingstehuizen	114
Bloedbanken	22
Consultatiebureaus voor Alcohol en Drugs	5
CPA's	28
Dagverblijf voor Ouderen	199
Dialyse-Centra	3
Dienstenstructuren (ANW-Diensten)	60
Dietisten	3386
Diverse Samenwerkingsverbanden	562
Ergotherapeuten	848
Eurotransplant	1
Fysiotherapeuten	11868
Gecombineerde Verpleeginrichtingen	910
Genezers	26322
Gezinsvervangend Tehuis	413
Gezondheidscentra	148
GGD	35
GGZ Instellingen (PUK/PAAZ)	53
Huisartsen	6196
Inrichting voor Auditief Gehandicapten	30
Inrichting voor Visueel Gehandicapten	8
Instellingen voor Jeugdtandverzorging	19
Kinderdagverblijf	101
Klinisch-Genetische Centra	9
Kraamcentra	520
Laboratoria(Huisartsenlab./Gemeensch.Lab/Gemeensch Apoth+Lab	74
Leveranciers Hulpmiddelen	8123
Logopedisten	3283

Zorgvestigingen

Zorgsoort	Aantal
Medisch Specialisten	1543
Mondhygienisten	975
Oefentherapeuten	2268
Optometristen	8
Overige Artsen	2872
Overige Instellingen	139
Pedicuren	11257
Podotherapeuten	871
Psychologische Zorgverleners	9527
R.I.A.G.G. Instelling	52
Radiotherapeutische Centra	6
RIBW	159
Schoonheidsspecialistes	2546
Sociaal Pedagogische Diensten	29
Tandarts-Specialisten (Dento-Maxillaire Orthopedie)	325
Tandarts-Specialisten (Mondziekten en Kaakchirurgie)	42
Tandartsen	7024
Tandheelkundige Centra	536
Tandtechnici/Tandprothetici	632
Thuiszorginstelling	1480
Trombosediensten	44
Verloskundigen	1120
Verpleeginrichtingen voor Psycho-Geriatrische Patienten	47
Verpleeginrichtingen voor Somatische Ziekten	36
Verpleegkundigen	79
Verzorgingshuizen	1516
Zelfstandige Behandelcentra	1101
Ziekenhuizen	514
Zwakzinnigen	311

Bijlage 2 Hoofdactiviteiten Handelsregister Zorgondernemingen

Hoofdactiviteit	Omschrijving	Aantal
2		3
10		2
111	Teelt van granen, peulvruchten en oliehoudende zaden	16
113	Teelt van groenten en wortel- en knolgewassen	26
118		1
119	Teelt van overige eenjarige gewassen	9
124	Teelt van pit- en steenvruchten	3
125	Teelt van overige boomvruchten, kleinfruit en noten	2
128	Teelt van specerijgewassen en van aromatische en medicinale gewassen	4
129	Teelt van overige meerjarige gewassen	1
13		2
130	Teelt van sierplanten	22
141	Fokken en houden van melkvee	163
142	Fokken en houden van runderen (geen melkvee)	29
143	Fokken en houden van paarden en ezels	15
145	Fokken en houden van schapen en geiten	42
146	Fokken en houden van varkens	17
147	Fokken en houden van pluimvee	6
149	Fokken en houden van overige dieren	13
15		1
150	Akker- en/of tuinbouw in combinatie met het fokken en houden van dieren	17
161	Dienstverlening voor de akker- en/of tuinbouw	34
162	Dienstverlening voor het fokken en houden van dieren	19
165		1
167		1
172		1
174		1
182		1
184		3
20		8
210	Bosbouw	1
226		1
228		3
229		3
23		4
24		2
252		2

Hoofdactiviteit	Omschrijving	Aantal
26		1
299		2
30		7
311	Zee- en kustvisserij	1
316		1
318		2
322	Kweken van zoetwatervis en -schaaldieren	1
33		5
343		2
344		1
345		2
346		1
35		4
36		2
40		2
412		1
416		1
45		2
475		1
478		1
485		1
492		1
497		1
499		1
50		2
511		1
515		1
518		1
524		1
53		1
541		1
562		1
566		1
570		2
573		1
575		1
58		1
592		1
594		1
598		1
599		1

Hoofdactiviteit	Omschrijving	Aantal
70		7
71		1
72		1
73		1
76		3
77		4
78		2
87		1
910	Dienstverlening voor de winning van aardolie en aardgas	1
1011	Slachterijen (geen pluimvee-)	1
1051	Vervaardiging van zuivelproducten (geen consumptie-ijs)	5
1052	Vervaardiging van consumptie-ijs	1
1061	Vervaardiging van meel (geen zetmeel)	3
1071	Vervaardiging van brood en vers banketbakkerswerk	10
1081	Vervaardiging van suiker	1
10822	Vervaardiging van chocolade en suikerwerk	1
1084	Vervaardiging van specerijen, sauzen en kruiderijen; zoutraffinage	2
1085	Vervaardiging van kant-en-klaarmaaltijden en snacks	1
1086	Vervaardiging van gehomogeniseerde voedingspreparaten en dieetvoeding	4
1089	Vervaardiging van overige voedingsmiddelen n.e.g.	7
1091	Vervaardiging van veevoeders	2
1092	Vervaardiging van voeders voor huisdieren	3
1102	Vervaardiging van wijn uit druiven	1
1320	Weven van textiel	2
1330	Textielveredeling	2
1391	Vervaardiging van gebreide en gehaakte stoffen	1
1392	Vervaardiging van geconfectioneerde artikelen van textiel (geen kleding)	6
1394	Vervaardiging van koord, bindgaren, touw en netten	1
1399	Vervaardiging van overige textielproducten n.e.g.	9
1412	Vervaardiging van werkkleding	2
1413	Vervaardiging van overige bovenkleding	16
1419	Vervaardiging van baby- en sportkleding en kledingaccessoires	1
1431	Vervaardiging van gebreide en gehaakte kousen en sokken	1
1512	Vervaardiging van koffers, tassen e.d. en van zadel- en tuigmakerswerk	5
1520	Vervaardiging van schoenen	3
16102	Verduurzamen van hout	1
16231	Vervaardiging van deuren, ramen en kozijnen van hout	1

Hoofdactiviteit	Omschrijving	Aantal
16239	Vervaardiging van overig timmerwerk voor de bouw (geen deuren, ramen en kozijnen)	5
1624	Vervaardiging van houten emballage	1
1629	Vervaardiging van overige artikelen van hout en van artikelen van kurk, riet en vlechtwerk (geen meubels)	4
17211	Vervaardiging van verpakkingsmiddelen van papier en karton	2
1722	Vervaardiging van huishoudelijke en sanitaire papierwaren	2
1729	Vervaardiging van overige papier- en kartonwaren	1
18122	Drukkerijen van tijdschriften	1
18123	Drukkerijen van reclame	2
18129	Overige drukkerijen n.e.g.	11
1813	Prepress- en premedia-activiteiten	2
1814	Grafische afwerking en overige activiteiten verwant aan de drukkerij	4
1820	Reproductie van opgenomen media	1
2011	Vervaardiging van industriële gassen	2
2013	Vervaardiging van overige anorganische basischemicaliën	5
20141	Vervaardiging van petrochemische producten	2
20149	Vervaardiging van overige organische basischemicaliën (geen petrochemische producten)	6
2015	Vervaardiging van kunstmeststoffen en stikstofverbindingen	2
2016	Vervaardiging van kunststof in primaire vorm	13
2020	Vervaardiging van verdelgingsmiddelen en overige landbouwchemicaliën	4
2041	Vervaardiging van zeep, wasmiddelen, poets- en reinigingsmiddelen	4
2042	Vervaardiging van parfums en cosmetica	22
2052	Vervaardiging van lijm en bereide kleefmiddelen	1
2059	Vervaardiging van overige chemische producten n.e.g.	9
2110	Vervaardiging van farmaceutische grondstoffen	47
2120	Vervaardiging van farmaceutische producten (geen grondstoffen)	322
2222	Vervaardiging van verpakkingsmiddelen van kunststof	3
2229	Vervaardiging van overige producten van kunststof	12
2312	Vormen en bewerken van vlakglas	5
2313	Vervaardiging van holglas	1
2319	Vervaardiging en bewerking van overig glas, inclusief technisch glaswerk	6
2320	Vervaardiging van vuurvaste keramische producten	1
2341	Vervaardiging van huishoudelijk en sieraardewerk	2
2370	Natuursteenbewerking	2
2442	Vervaardiging van aluminium	2
2511	Vervaardiging van metalen constructiewerken en delen daarvan	10

Hoofdactiviteit	Omschrijving	Aantal
2550	Smeden, persen, stampen en profielwalsen van metaal; poedermetallurgie	1
2561	Oppervlaktebehandeling en bekleding van metaal	2
2562	Algemene metaalbewerking	24
2572	Vervaardiging van hang- en sluitwerk	1
2573	Vervaardiging van gereedschap	4
2599	Vervaardiging van overige producten van metaal n.e.g.	4
2611	Vervaardiging van elektronische componenten	15
2612	Vervaardiging van elektronische printplaten	2
2620	Vervaardiging van computers en randapparatuur	2
2630	Vervaardiging van communicatieapparatuur	7
2640	Vervaardiging van consumentenelektronica	5
2651	Vervaardiging van meet-, regel-, navigatie- en controleapparatuur	583
2660	Vervaardiging van bestralingsapparatuur en van elektromedische en elektrotherapeutische apparatuur	136
2670	Vervaardiging van optische instrumenten en apparatuur	166
2680	Vervaardiging van informatiedragers	1
2711	Vervaardiging van elektromotoren, elektrische generatoren en transformatoren	10
2712	Vervaardiging van schakel- en verdeelinrichtingen	5
2720	Vervaardiging van batterijen en accumulatoren	3
2732	Vervaardiging van overige elektrische en elektronische kabel	2
2740	Vervaardiging van elektrische lampen en verlichtingsapparaten	4
2751	Vervaardiging van elektrische huishoudapparaten	7
2752	Vervaardiging van niet-elektrische huishoudapparaten	2
2790	Vervaardiging van overige elektrische apparatuur	6
2813	Vervaardiging van niet-hydraulische pompen en compressoren	2
2814	Vervaardiging van appendages	1
2815	Vervaardiging van tandwielen, lagers en overige drijfwerkelementen	4
2821	Vervaardiging van industriële ovens en branders	1
2822	Vervaardiging van hijs-, hef- en transportwerktuigen	4
2824	Vervaardiging van pneumatisch en elektrisch handgereedschap	1
2825	Vervaardiging van machines en apparaten voor industriële koeltechniek en klimaatregeling	3
2829	Vervaardiging van overige machines en apparaten voor algemeen gebruik n.e.g.	9
2830	Vervaardiging van machines en werktuigen voor de land- en bosbouw	2
2841	Vervaardiging van gereedschapswerktuigen voor metaalbewerking	2
2849	Vervaardiging van gereedschapswerktuigen niet voor metaalbewerking	1

Hoofdactiviteit	Omschrijving	Aantal
2891	Vervaardiging van machines voor de metallurgie	1
2892	Vervaardiging van machines voor de bouw en winning van delfstoffen	1
2893	Vervaardiging van machines voor de productie van voedings- en genotmiddelen	1
2899	Vervaardiging van overige machines, apparaten en werktuigen voor specifieke doeleinden n.e.g.	9
29201	Carrosseriebouw	2
2932	Vervaardiging van niet-elektrische en -elektronische onderdelen en toebehoren voor auto's	1
3		86
3011	Bouw van schepen en drijvend materieel (geen sport- en recreatievaartuigen)	3
3012	Bouw van sport- en recreatievaartuigen	5
3020	Vervaardiging van rollend spoor- en tramwegmaterieel	1
3030	Vervaardiging van vliegtuigen en onderdelen daarvoor	3
3092	Vervaardiging van fietsen en invalidenwagens	15
3099	Vervaardiging van overige transportmiddelen n.e.g.	2
31011	Interieurbouw	8
31012	Vervaardiging van bedrijfsmeubels	4
3102	Vervaardiging van keukenmeubels	1
3103	Vervaardiging van matrassen	1
3109	Vervaardiging van overige meubels	17
3212	Bewerken van edelstenen en vervaardiging van sieraden e.d. (geen imitatie)	14
3213	Vervaardiging van imitatiesieraden	4
3220	Vervaardiging van muziekinstrumenten	7
3230	Vervaardiging van sportartikelen	3
3240	Vervaardiging van speelgoed en spellen	3
32501	Tandtechnische bedrijven	1515
32502	Vervaardiging van medische instrumenten en hulpmiddelen (geen tandtechniek)	987
325029		22
32991	Sociale werkvoorziening	64
32999	Vervaardiging van overige goederen n.e.g.	7
33121	Reparatie en onderhoud van machines voor algemeen gebruik en machine-onderdelen (geen gereedschap)	11
33122	Reparatie en onderhoud van pneumatisch en elektrisch gereedschap en gereedschapswerktuigen	2
33123	Reparatie en onderhoud van machines voor een specifieke bedrijfstak	9
3313	Reparatie van elektronische en optische apparatuur	117
331302		1

Hoofdactiviteit	Omschrijving	Aantal
3314	Reparatie van elektrische apparatuur	267
3315	Reparatie en onderhoud van schepen	8
3316	Reparatie en onderhoud van vliegtuigen	1
3317	Reparatie en onderhoud van overige transportmiddelen	1
3319	Reparatie van overige apparatuur	45
33221	Installatie van machines voor algemeen gebruik (geen gereedschap)	9
33223	Installatie van machines voor een specifieke bedrijfstak	3
3323	Installatie van elektronische en optische apparatuur	266
3324	Installatie van elektrische apparatuur	352
3329	Installatie van overige toebehoren	29
35111	Productie van elektriciteit door thermische, kern- en warmtekrachtcentrales	2
35112	Productie van elektriciteit door windenergie	3
35113	Productie van elektriciteit door zonnecellen, warmtepompen en waterkracht	4
3513	Distributie van elektriciteit en gasvormige brandstoffen via leidingen	4
3514	Handel in elektriciteit en in gas via leidingen	1
3600	Winning en distributie van water	3
3700	Afvalwaterinzameling en -behandeling	2
3811	Inzameling van onschadelijk afval	12
3821	Behandeling van onschadelijk afval	2
3822	Behandeling van schadelijk afval	2
3832	Gesorteerd materiaal voorbereiden tot recycling	3
3900	Sanering en overig afvalbeheer	2
4110	Projectontwikkeling	18
4120	Algemene burgerlijke en utiliteitsbouw	188
42112	Stratenmaken	4
4221	Leggen van rioleringen, buizen en pijpleidingen; aanleg van bronbemaling	1
4222	Leggen van elektriciteits- en telecommunicatiekabels	11
4311	Slopen van bouwwerken	1
4312	Grondverzet	10
4321	Elektrotechnische bouwinstallatie	102
43221	Loodgieters- en fitterswerk; installatie van sanitair	18
43222	Installatie van verwarmings- en luchtbehandelingsapparatuur	12
4329	Overige bouwinstallatie	13
4331	Stukadoren	3
4332	Bouwtimmeren	29
4333	Afwerking van vloeren en wanden	11
4334	Schilderen en glaszetten	19

Hoofdactiviteit	Omschrijving	Aantal
4339	Overige afwerking van gebouwen	3
43992	Vlechten van betonstaal	2
43993	Metselen en voegen	5
43999	Overige gespecialiseerde werkzaamheden in de bouw n.e.g.	15
45111	Import van nieuwe personenauto's en lichte bedrijfsauto's	1
45112	Handel in en reparatie van personenauto's en lichte bedrijfsauto's (geen import van nieuwe)	38
45192	Handel in en reparatie van zwaardere bedrijfsauto's (geen import van nieuwe)	1
45193	Handel in en reparatie van aanhangwagens en opleggers (geen caravans)	1
45194	Handel in en reparatie van caravans	2
45202	Bandenservicebedrijven	1
45203	Reparatie van specifieke auto-onderdelen	4
45204	Carrosserieherstel	2
45205	Overige gespecialiseerde reparatie; wassen en slepen van auto's	7
45311	Groothandel en handelsbemiddeling in auto-onderdelen en -accessoires (geen banden)	4
4532	Detailhandel in auto-onderdelen en -accessoires	1
45401	Groothandel en handelsbemiddeling in motorfietsen en onderdelen daarvan	3
45402	Detailhandel in en reparatie van motorfietsen en onderdelen daarvan	5
4611	Handelsbemiddeling in landbouwproducten, levende dieren en grondstoffen voor textiel en voedingsmiddelen	7
4612	Handelsbemiddeling in brandstoffen, ertsen, metalen en chemische producten	3
4613	Handelsbemiddeling in hout, vlakglas, sanitair en bouwmaterialen	1
4614	Handelsbemiddeling in machines, technische benodigdheden, schepen en vliegtuigen	18
4615	Handelsbemiddeling in meubels, huishoudelijke artikelen en ijzerwaren	6
4616	Handelsbemiddeling in textiel, kleding, schoeisel en lederwaren	7
4617	Handelsbemiddeling in voedings- en genotmiddelen	35
4618	Handelsbemiddeling gespecialiseerd in overige goederen	82
4619	Niet-gespecialiseerde handelsbemiddeling	13
46211	Groothandel in granen	2
46213	Groothandel in hooi, stro en ruwvoeder	1
46214	Groothandel in meng- en krachtvoeder	3
46218	Groothandel in akkerbouwproducten en veevoeder algemeen assortiment	9
46219	Groothandel in overige akkerbouwproducten	3
4622	Groothandel in bloemen en planten	1

Hoofdactiviteit	Omschrijving	Aantal
46231	Groothandel in levend vee	9
46232	Groothandel in huisdieren, siervissen, siervogels en wilde dieren	2
46241	Groothandel in huiden en vellen	1
46311	Groothandel in groenten en fruit	1
4632	Groothandel in vlees en vleeswaren en in wild en gevogelte (niet levend)	3
46331	Groothandel in zuivelproducten en spijsoliën en -vetten	2
46332	Groothandel in eieren	1
4634	Groothandel in dranken (geen zuivel)	7
4635	Groothandel in tabaksproducten en rokersbenodigdheden	1
4636	Groothandel in suiker, chocolade en suikerwerk	1
4637	Groothandel in koffie, thee, cacao en specerijen (geen ruwe)	2
46381	Groothandel in snacks	1
46383	Gespecialiseerde groothandel in overige voedings- en genotmiddelen n.e.g.	70
46384	Groothandel in bakkerijgrondstoffen	1
46389	Gespecialiseerde groothandel in overige grondstoffen en halffabrikaten voor de voedingsmiddelenindustrie n.e.g.	4
4639	Groothandel in voedings- en genotmiddelen algemeen assortiment	82
46412	Groothandel in huishoudtextiel en beddengoed	4
46421	Groothandel in bovenkleding	15
46423	Groothandel in onderkleding	2
46424	Groothandel in schoenen	4
46425	Groothandel in mode-artikelen	6
46429	Groothandel in textielwaren algemeen assortiment	2
46431	Groothandel in witgoed	1
46432	Groothandel in elektrische huishoudelijke apparatuur (geen witgoed, audio en video)	6
46433	Groothandel in audio- en video-apparatuur	9
46434	Groothandel in fotografische artikelen	14
46435	Groothandel in beeld- en geluidsdragers	2
46436	Groothandel in optische artikelen	31
46441	Groothandel in glas, porselein en aardewerk	5
46442	Groothandel in was-, poets- en reinigingsmiddelen	13
4645	Groothandel in parfums en cosmetica	84
46461	Groothandel in farmaceutische producten	1035
46462	Groothandel in medische en tandheelkundige instrumenten, verpleeg- en orthopedische artikelen en laboratoriumbenodigdheden	2243
46471	Groothandel in huismeubilair	15
46472	Groothandel in woningtextiel en vloerbedekking	3
46473	Groothandel in verlichtingsartikelen	10

Hoofdactiviteit	Omschrijving	Aantal
4648	Groothandel in juweliersartikelen en uurwerken	6
46491	Groothandel in fietsen en bromfietsen	3
46492	Groothandel in watersportartikelen	3
46493	Groothandel in kampeerartikelen (geen caravans)	1
46494	Groothandel in speelgoed	4
46495	Groothandel in muziekinstrumenten	2
46496	Groothandel in sportartikelen (geen watersport)	20
46497	Groothandel in papier- en kartonwaren (geen verpakkingsmateriaal)	3
46498	Groothandel in boeken, tijdschriften en overig drukwerk	8
46499	Groothandel in overige consumentenartikelen (non-food) n.e.g.	39
4651	Groothandel in computers, randapparatuur en software	26
4652	Groothandel in elektronische en telecommunicatieapparatuur en bijbehorende onderdelen	45
4661	Groothandel in landbouwmachines, werktuigen en tractoren	6
4662	Groothandel in gereedschapswerktuigen	3
4665	Groothandel in bedrijfsmeubels	8
4666	Groothandel in kantoormachines (geen computers)	1
46681	Groothandel in machines voor de grafische industrie	3
46682	Groothandel in machines voor de voedings- en genotmiddelenindustrie	3
46691	Groothandel in intern transportmaterieel	1
46692	Groothandel in machines en apparaten voor de warmte-, koel- en vriestechniek	3
46693	Groothandel in verbrandingsmotoren, pompen en compressoren	2
46694	Groothandel in appendages, technische toebehoren e.d.	7
46695	Groothandel in meet- en regelapparaten	72
46696	Groothandel in emballage	7
46697	Groothandel in scheepsbenodigdheden en visserij-artikelen	10
46699	Groothandel in overige machines, apparaten en toebehoren voor industrie en handel n.e.g.	36
46711	Groothandel in vaste brandstoffen	1
46712	Groothandel in vloeibare en gasvormige brandstoffen	2
46713	Groothandel in minerale olieproducten (geen brandstoffen)	3
46721	Groothandel in metaalertsen	3
46722	Groothandel in ferrometalen en -halffabrikaten	2
46723	Groothandel in non-ferrometalen en -halffabrikaten	1
46731	Groothandel in hout en plaatmateriaal	2
46732	Groothandel in verf en verfwaren	1
46736	Groothandel in tegels en plavuizen	1
46737	Groothandel in sanitaire artikelen en sanitair installatiemateriaal	5
46738	Groothandel gespecialiseerd in overige bouwmaterialen	3

Hoofdactiviteit	Omschrijving	Aantal
46739	Groothandel in bouwmaterialen algemeen assortiment	3
46741	Groothandel in ijzer- en metaalwaren	4
46751	Groothandel in chemische grondstoffen en chemicaliën voor industriële toepassing	64
46752	Groothandel in bestrijdingsmiddelen en kunstmeststoffen	7
46761	Groothandel in textielgrondstoffen en -halffabrikaten	2
46772	Groothandel in ijzer- en staalschroot en oude non-ferrometalen	1
46779	Groothandel in overige oude materialen en afvalstoffen	8
46901	Niet-gespecialiseerde groothandel in consumentenartikelen	21
46902	Niet-gespecialiseerde groothandel in niet-consumentenartikelen	8
4711	Supermarkten en dergelijke winkels met een algemeen assortiment voedings- en genotmiddelen	740
47191	Warenhuizen	9
47192	Winkels met een algemeen assortiment non-food (geen warenhuizen)	7
4721	Winkels in aardappelen, groenten en fruit	3
47221	Winkels in vlees en vleeswaren	4
4723	Winkels in vis	2
47241	Winkels in brood en banket	4
47242	Winkels in chocolade en suikerwerk	2
4725	Winkels in dranken	8
4726	Winkels in tabaksproducten	9
47292	Winkels in natuurvoeding en reformartikelen	96
47293	Winkels in buitenlandse voedingsmiddelen	8
47299	Gespecialiseerde winkels in overige voedings- en genotmiddelen n.e.g.	24
4730	Benzinestations	1
4741	Winkels in computers, randapparatuur en software	4
4742	Winkels in telecommunicatieapparatuur	2
47431	Winkels in audio- en videoapparatuur	1
47432	Winkels in een algemeen assortiment van wit- en bruingoed	3
47512	Winkels in huishoudtextiel	1
47513	Winkels in breiwol, handwerken en fournituren	1
47521	Winkels in ijzerwaren en gereedschappen	2
47522	Winkels in verf, verfwaren en behang	10
47523	Winkels in houten bouw- en tuinmaterialen	1
47527	Winkels gespecialiseerd in overige doe-het-zelfartikelen	1
4753	Winkels in vloerbedekking en gordijnen	1
47543	Winkels in overige elektrische huishoudelijke apparatuur	1
47544	Winkels in onderdelen voor elektrische huishoudelijke apparatuur	1
47591	Winkels in meubels	6

Hoofdactiviteit	Omschrijving	Aantal
47592	Winkels in verlichtingsartikelen	4
47593	Winkels in artikelen voor woninginrichting algemeen assortiment	6
47594	Winkels in muziekinstrumenten	1
47596	Winkels gespecialiseerd in overige huishoudelijke artikelen n.e.g.	5
47597	Winkels in huishoudelijke artikelen algemeen assortiment	10
4761	Winkels in boeken	15
4762	Winkels in kranten, tijdschriften en kantoorbehoeften	7
47641	Winkels in fietsen en bromfietsen	8
47642	Winkels in watersportartikelen	1
47643	Winkels in sportartikelen (geen watersport)	8
4765	Winkels in speelgoed	12
47711	Winkels in herenkleding	1
47712	Winkels in dameskleding	8
47713	Winkels in bovenkleding en mode-artikelen (algemeen assortiment)	16
47714	Winkels in baby- en kinderkleding	3
47715	Winkels in babyartikelen algemeen assortiment	6
47716	Winkels in onderkleding, foundations e.d.	14
47717	Winkels in mode-artikelen	7
47718	Textielsupermarkten	1
47721	Winkels in schoenen	33
47722	Winkels in lederwaren en reisartikelen	1
4773	Apotheken	2516
47741	Winkels in drogisterij-artikelen	2925
47742	Winkels in medische en orthopedische artikelen	1075
4775	Winkels in parfums en cosmetica	49
47761	Winkels in bloemen en planten, zaden en tuinbenodigdheden	11
47762	Tuincentra	3
47763	Winkels in dieren, dierbenodigdheden en hengelsportartikelen	17
4777	Winkels in juweliersartikelen en uurwerken	9
47782	Winkels in optische artikelen	460
47783	Winkels in schilderijen, lijsten, prenten, kunstvoorwerpen en religieuze artikelen	12
47789	Winkels gespecialiseerd in overige artikelen n.e.g.	61
47791	Winkels in antiek	6
47792	Winkels in tweedehands kleding	7
47793	Winkels in tweedehands goederen (geen kleding)	28
47811	Markthandel in aardappelen, groenten en fruit	1
47819	Markthandel in overige voedings- en genotmiddelen	16
4782	Markthandel in textiel, kleding en schoenen	4
47892	Markthandel in tweedehands goederen	6

Hoofdactiviteit	Omschrijving	Aantal
47899	Markthandel in overige goederen	18
47911	Detailhandel via internet in voedingsmiddelen en drogisterijwaren	52
47912	Detailhandel via internet in consumentenelektronica	10
47913	Detailhandel via internet in boeken, tijdschriften, cd's, dvd's	6
47914	Detailhandel via internet in kleding en mode-artikelen	37
47915	Detailhandel via internet in huis- en tuinartikelen	34
47916	Detailhandel via internet in vrijetijdsartikelen	18
47918	Detailhandel via internet in overige non-food	76
47919	Detailhandel via internet in een algemeen assortiment non-food	26
47991	Colportage	9
47992	Straathandel	2
47999	Detailhandel via overige distributievormen	69
4932	Vervoer per taxi	22
49391	Ongeregeld personenvervoer over de weg	7
49393	Geregeld besloten personenvervoer over de weg	3
4941	Goederenvervoer over de weg (geen verhuizingen)	41
4942	Verhuisvervoer	2
5010	Zee- en kustvaart (passagiersvaart en veerdiensten)	2
5030	Binnenvaart (passagiersvaart en veerdiensten)	7
50401	Binnenvaart (vrachtvaart)	1
50403	Binnenvaart (sleep- en duwvaart)	3
5110	Personenvervoer door de lucht	4
52102	Opslag in koelhuizen e.d.	2
52109	Opslag in distributiecentra en overige opslag (niet in tanks, koelhuizen e.d.)	68
5221	Dienstverlening voor vervoer over land	79
5222	Dienstverlening voor vervoer over water	4
5223	Dienstverlening voor de luchtvaart	2
52242	Laad-, los- en overslagactiviteiten niet voor zeevaart	3
52291	Expediteurs, cargadoors, bevrachters en andere tussenpersonen in het goederenvervoer	39
52292	Weging en meting	3
53201	Post zonder universele dienstverplichting	2
53202	Koeriers	34
55101	Hotel-restaurants	11
55102	Hotels (geen hotel-restaurants), pensions en conferentie-oorden	35
55201	Verhuur van vakantiehuisjes en appartementen	14
55202	Jeugdherbergen en vakantiekampen	1
5530	Kampeerterreinen	31
5590	Overige logiesverstrekking	6
56101	Restaurants	45

Hoofdactiviteit	Omschrijving	Aantal
56102	Cafetaria's, lunchrooms, snackbars, ijssalons, eetkramen e.d.	230
5621	Eventcatering	71
5629	Kantines en contractcatering	16
5630	Cafés	94
5811	Uitgeverijen van boeken en databanken	53
5813	Uitgeverijen van kranten	1
5814	Uitgeverijen van tijdschriften	13
5819	Overige uitgeverijen (niet van software)	4
59111	Productie van films (geen televisiefilms)	26
59112	Productie van televisieprogramma's	5
5912	Facilitaire activiteiten voor film- en televisieproductie	17
5914	Bioscopen	1
5920	Maken en uitgeven van geluidsopnamen	9
6110	Draadgebonden telecommunicatie	1
6190	Overige telecommunicatie	10
6201	Ontwikkelen, produceren en uitgeven van software	162
6202	Advisering op het gebied van informatietechnologie	96
6203	Beheer van computerfaciliteiten	7
6209	Overige dienstverlenende activiteiten op het gebied van informatietechnologie	39
6311	Gegevensverwerking, webhosting en aanverwante activiteiten	17
6312	Webportals	27
6391	Persagentschappen	12
6399	Overige dienstverlenende activiteiten op het gebied van informatie n.e.g.	2
6420	Financiële holdings	627
64301	Beleggingsinstellingen in financiële activa	15
64302	Beleggingsinstellingen in vaste activa	8
64303	Beleggingsinstellingen met beperkte toetreding	90
64922	Volkskredietbanken en commerciële financieringsmaatschappijen	3
64923	Participatiemaatschappijen	24
6499	Overige financiële intermediatie	3
65111	Levensverzekeringen (geen herverzekering)	15
65112	Naturaverzekeringen	1
65113	Spaarkassen	1
6512	Schadeverzekeringen (geen herverzekering)	800
6520	Herverzekering	2
65301	Bedrijfspensioenfondsen	4
65302	Ondernemingspensioenfondsen en -spaarfondsen	12
65309	Overige pensioenfondsen	1
6612	Commissionairs en makelaars in effecten, beleggingsadviseurs e.d.	11

Hoofdactiviteit	Omschrijving	Aantal
66191	Administratiekantoren voor aandelen en obligaties	32
66193	Hypotheek- en kredietbemiddeling, geldwisselkantoren, bank- en spaaragentschappen e.d.	29
6621	Risicoanalisten en schadetaxateurs	14
6622	Assurantietussenpersonen	80
66292	Actuariële en pensioenadviesbureaus; beheer en administratie van pensioenvermogens	8
66293	Waarborgfondsen	2
66299	Overige dienstverlening op het gebied van verzekeringen en pensioenfondsen n.e.g.	2
6630	Vermogensbeheer	12
6810	Handel in eigen onroerend goed	26
68201	Woningbouwverenigingen en -stichtingen	6
68203	Verhuur van overige woonruimte	36
68204	Verhuur van onroerend goed (niet van woonruimte)	60
6831	Bemiddeling bij handel, huur of verhuur van onroerend goed	33
6832	Beheer van onroerend goed	40
69101	Advocatenkantoren	29
69102	Rechtskundige adviesbureaus	150
69104	Deurwaarderskantoren	2
69105	Octrooibureaus	7
69202	Accountants-administratieconsulenten	14
69203	Boekhoudkantoren	172
69204	Belastingconsulenten	20
69209	Overige administratiekantoren	195
70101	Concerndiensten binnen eigen concern	25
70102	Holdings (geen financiële)	146
7021	Public relationsbureaus	27
70221	Organisatie-adviesbureaus	2104
70222	Advisering op het gebied van management en bedrijfsvoering (geen public relations en organisatieadviesbureaus)	196
7111		16
71111	Architecten (geen interieurarchitecten)	1
7112	Ingenieurs en overig technisch ontwerp en advies	316
71201	Keuring en controle van agrarische producten en voedingsmiddelen	1
71202	Keuring en controle van machines, apparaten en materialen	30
71203	Overige keuring en controle	5
72111	Biotechnologisch speur- en ontwikkelingswerk op het gebied van agrarische producten en processen	4
72112	Biotechnologisch speur- en ontwikkelingswerk op het gebied van medische producten en farmaceutische processen en van voeding	55

Hoofdactiviteit	Omschrijving	Aantal
72113	Biotechnologisch speur- en ontwikkelingswerk voor overige toepassingen	5
72191	Speur- en ontwikkelingswerk op het gebied van landbouw en visserij (niet biotechnologisch)	21
72192	Technisch speur- en ontwikkelingswerk	43
72193	Speur- en ontwikkelingswerk op het gebied van gezondheid en voeding (niet biotechnologisch)	1925
72199	Overig natuurwetenschappelijk speur- en ontwikkelingswerk (niet biotechnologisch)	11
7220	Speur- en ontwikkelingswerk op het gebied van de maatschappij- en geesteswetenschappen	48
7311	Reclamebureaus	162
7312	Handel in advertentieruimte en -tijd	6
7320	Markt- en opinieonderzoekbureaus	78
7410		9
74101	Communicatie- en grafisch ontwerp	22
74102	Industrieel en productontwerp	16
74103	Interieur- en ruimtelijk ontwerp	23
74201	Fotografie	123
7430	Vertalers en tolken	93
7490	Overige specialistische zakelijke dienstverlening	237
7500	Veterinaire dienstverlening	2570
77112	Operational lease van personenauto's en lichte bedrijfsauto's	2
7712	Verhuur en lease van vrachtwagens, autobussen, caravans en aanhangwagens	1
7721	Verhuur van sport- en recreatieartikelen	9
7722	Videotheken	2
77292	Verhuur van kleding en huishoudelijke artikelen	3
77299	Verhuur van overige consumentenartikelen n.e.g.	10
7732	Verhuur en lease van machines en installaties voor de bouw	1
77399	Verhuur en lease van overige machines en werktuigen en van overige goederen (geen automaten)	13
7740	Lease van niet-financiële immateriële activa	29
7810	Arbeidsbemiddeling	505
78201	Uitzendbureaus	62
78202	Uitleenbureaus	33
78203	Banenpools (werkgelegenheidsprojecten)	7
7830	Payrolling (personeelsbeheer)	2
7911	Reisbemiddeling	182
7912	Reisorganisatie	35
7990	Informatieverstrekking op het gebied van toerisme en reserveringsbureaus	71

Hoofdactiviteit	Omschrijving	Aantal
8010	Particuliere beveiliging	34
8020	Beveiliging via beveiligingssystemen	5
8030	Opsporing	4
8110	Facility management	44
8121	Interieurreiniging van gebouwen	125
8122		3
81221	Glazenwassen	12
81222	Schoorsteenvegen	1
81229	Overige gespecialiseerde reiniging van gebouwen en industriële reiniging.	6
8129	Overige reiniging	12
8130	Landschapsverzorging	58
8211	Brede administratieve dienstverlening	114
8219	Secretariële dienstverlening	72
8220	Callcenters	5
8230	Organiseren van congressen en beurzen	15
8291	Kredietinformatie- en incassobureaus	7
8292	Pakken, sorteren e.d. in loon	11
82999	Overige zakelijke dienstverlening n.e.g.	52
8411	Algemeen overheidsbestuur	1768
8412	Openbaar bestuur op het gebied van gezondheidszorg, onderwijs, cultuur en welzijn (geen sociale verzekeringen)	437
8413	Openbaar bestuur op het gebied van het bedrijfsleven	1
84231	Rechtspraak	1
84232	Ministerie van Justitie en gevangeniswezen	2
8424	Politie	2
8425	Brandweer	64
8430	Verplichte sociale verzekeringen	2
85201	Basisonderwijs voor leerplichtigen	217
85202	Speciaal basisonderwijs	64
85203	Speciaal onderwijs in expertisecentra	28
85311	Havo en vwo	12
85312	Voorbereidend middelbaar beroepsonderwijs	35
85313	Praktijkonderwijs	8
85314	Brede scholengemeenschappen voor voortgezet onderwijs	118
85321	Middelbaar beroepsonderwijs	30
85322	Educatie	72
85323	Middelbaar beroepsonderwijs en educatie (gecombineerd)	32
8541	Niet-universitair hoger onderwijs	34
8542	Universitair hoger onderwijs	41
85511	Zeil- en surfscholen	6

Hoofdactiviteit	Omschrijving	Aantal
85519	Overig sport- en recreatieonderwijs	783
85521	Dansscholen	41
85522	Kunstzinnige vorming van amateurs (geen dansscholen)	180
8553	Auto- en motorrijscholen	29
85591	Afstandsonderwijs	8
85592	Bedrijfsopleiding en -training	1533
85599	Studiebegeleiding, vorming en onderwijs n.e.g.	729
8560	Dienstverlening voor het onderwijs	2074
86101	Universitair medische centra	31
86102	Algemene ziekenhuizen	246
86103	Categorale ziekenhuizen	260
86104	Geestelijke gezondheids- en verslavingszorg met overnachting	1322
8621	Praktijken van huisartsen	9515
86221	Praktijken van medisch specialisten en medische dagbehandelcentra (geen tandheelkunde of psychiatrie)	7312
862211		1
86222	Praktijken van psychiaters en dagbehandelcentra voor geestelijke gezondheids- en verslavingszorg	1636
86231	Praktijken van tandartsen	9049
86232	Praktijken van tandheelkundig specialisten	715
86911	Praktijken van verloskundigen	1683
86912	Praktijken van fysiotherapeuten	10143
86913	Praktijken van psychotherapeuten en psychologen	15779
86919	Overige paramedische praktijken (geen fysiotherapie en psychologie) en alternatieve genezers	33265
869191		1
869199		11
86921	Gezondheidscentra	1107
86922	Arbobegeleiding en reïntegratie	3222
86923	Preventieve gezondheidszorg (geen arbobegeleiding)	2975
86924	Medische laboratoria, trombodediensten en overig behandelingsondersteunend onderzoek	555
86925	Ambulancediensten en centrale posten	163
86929	Samenwerkingsorganen op het gebied van gezondheidszorg en overige gezondheidszorgondersteunende diensten	6685
869291		1
8710	Verpleeghuizen	872
8720	Huizen en dagverblijven voor verstandelijk gehandicapten en psychiatrische cliënten	6527
872001		25
87301	Huizen en dagverblijven voor niet-verstandelijk gehandicapten	420
873011		50

Hoofdactiviteit	Omschrijving	Aantal
87302	Verzorgingshuizen	1829
87901	Jeugdzorg met overnachting en dagverblijven voor jeugdzorg	1031
87902	Maatschappelijk opvang met overnachting	627
879021		1
88101	Thuiszorg	13751
88102	Welzijnswerk voor ouderen	1135
88103	Ondersteuning en begeleiding van gehandicapten	2270
88911	Kinderopvang	15922
88912	Peuterspeelzaalwerk	1986
88991	Ambulante jeugdzorg	1376
88992	Maatschappelijk werk	5256
889922		1
88993	Lokaal welzijnswerk	18367
889931		4
889932		1
88999	Overig maatschappelijk advies, gemeenschapshuizen en samenwerkingsorganen op het gebied van welzijn	18027
889993		4
889994		1
889999		12
90011	Beoefening van podiumkunst	208
90012	Producenten van podiumkunst	25
90013	Circus en variété	20
9002	Dienstverlening voor uitvoerende kunst	147
9003	Schrijven en overige scheppende kunst	393
90041	Theaters en schouwburgen	6
91011	Openbare bibliotheken	31
91012	Kunstuitleencentra	2
91019	Overige culturele uitleencentra en openbare archieven	11
91021	Musea	18
91022	Kunstgalerieën en -expositieruimten	4
9103	Monumentenzorg	7
91041	Dieren- en plantentuinen, kinderboerderijen	45
91042	Natuurbehoud	12
92001	Loterijen en kansspelen (geen amusementsautomaten)	1
92009	Exploitatie van amusements- en speelautomaten	6
93111	Zwembaden	25
93112	Sporthallen, sportzalen en gymzalen	202
93113	Sportvelden	3
93119	Overige sportaccommodaties	8
93121	Veldvoetbal	2

Hoofdactiviteit	Omschrijving	Aantal
93124	Tennis	1
93125	Paardensport en maneges	38
93129	Overige buitensport	4
9313	Fitnesscentra	144
93141	Individuele zaalsport	2
93143	Kracht- en vechtsport	4
93146	Sportscholen	66
93149	Overige binnensport en omnisport	1
93151	Zwem- en onderwatersport	4
93152	Roei-, kano-, zeil- en surfsport e.d.	1
93191	Beroepssportlieden	2
93192	Hengelsport	1
93194	Supportersverenigingen (sport)	1
93195	Organiseren van sportevenementen	34
93196	Overkoepelende organen en samenwerkings- en adviesorganen op het gebied van sport	16
93199	Overige sportactiviteiten n.e.g.	18
93211	Pret- en themaparken	5
93291	Jachthavens	3
93299	Overige recreatie n.e.g.(geen jachthavens)	183
9411	Bedrijfs- en werkgeversorganisaties	21
9412	Beroepsorganisaties	38
9420	Werknemersorganisaties	2
94911	Religieuze organisaties	53
94919	Overige levensbeschouwelijke organisaties	17
9492	Politieke organisaties	1
94991	Gezelligheidsverenigingen	40
94992	Hobbyclubs	10
94993	Steunfondsen (niet op het gebied van welzijnszorg)	199
94994	Vriendenkringen op het gebied van cultuur, fanclubs en overige kunstbevordering	13
94995	Overkoepelende organen en samenwerkings- en adviesorganen (niet op het gebied van gezondheids- en welzijnszorg, sport en recreatie)	22
94996	Overige ideële organisaties n.e.g.	44
94997	Overige belangenbehartiging n.e.g.	169
9511	Reparatie van computers en randapparatuur	7
9512	Reparatie van communicatieapparatuur	1
9521	Reparatie van consumentenelektronica (geen computers)	5
9522	Reparatie van elektrische huishoudelijke apparaten	9
9523	Reparatie van schoenen en lederwaren	18

Hoofdactiviteit	Omschrijving	Aantal
9524	Reparatie en stoffering van meubels	16
9525	Reparatie van uurwerken en juweliersartikelen	2
9529	Reparatie van overige consumentenartikelen	16
96011	Wasserijen en linnenverhuur	3
96012	Chemische wasserijen en ververijen	3
96021	Haarverzorging	144
96022	Schoonheidsverzorging, visagie, pedicures en manicures	776
96031	Uitvaartverzorging	27
96032	Crematoria, mortuaria en begraafplaatsen	18
9604	Sauna's, solaria, baden e.d.	284
9609	Overige dienstverlening n.e.g.	185
9700	Huishoudens als werkgever van huishoudelijk personeel	3
9900	Extraterritoriale organisaties en lichamen	2
		494

Bijlage 3 Beroepsgroepen BIG

Beroepsgroep	Omschrijving	Aantal	Type *)
17	Apothekers	6676	Art.3
1	Artsen	79698	Art.3
89	Diëtisten	4296	Art.34
90	Ergotherapeuten	4575	Art.34
4	Fysiotherapeuten	32502	Art.3
25	Gz-psychologen	16002	Art.3
88	Huidtherapeuten	616	Art.34
91	Logopedisten	7037	Art.34
92	Mondhygiënisten	1468	Art.34
94	Oefentherapeuten Cesar	1376	Art.34
93	Oefentherapeuten Mensendieck	1118	Art.34
87	Optometristen	681	Art.34
95	Orthoptisten	448	Art.34
96	Podotherapeuten	678	Art.34
16	Psychotherapeuten	6739	Art.3
97	Radiodiagnostisch laboranten	3787	Art.34
98	Radiotherapeutisch laboranten	1389	Art.34
2	Tandartsen	13944	Art.3
3	Verloskundigen	3727	Art.3
30	Verpleegkundigen	178621 7	Art.3

*) Type-indeling op basis van informatie op de website www.bigregister.nl

Art. 3 beroepen	Wettelijk beschermde beroepstitel BIG-register Tuchtrecht
Art. 34 beroepen	Wettelijk beschermde opleidingstitel Geen wettelijk register Geen tuchtrecht

Bijlage 4 Datakwaliteit

1 Data Kwaliteit Analyse

In een schoning wordt de kwaliteit van de data beoordeeld en gecorrigeerd voor wat betreft de volgende aspecten:

- De kwaliteit van tenaamstellingen.
- De kwaliteit van adressen.
- De kwaliteit van emailadressen.
- De kwaliteit van telefoonnummers.

1.1 Kwaliteitsbepaling van tenaamstellingen

De kwaliteitsbepaling voor tenaamstellingen wordt uitgevoerd met behulp van de datakwaliteitscomponent HIquality Name. Deze module bevat een aantal functionaliteiten dat gebruikt wordt om een tenaamstelling te controleren en waar mogelijk aan te passen:

1) Formaatconversie:
    - Bedrijf- en persoonherkenning.
    - Herkenning van individuele naamelementen.
    - Uniformering van naamelementen (interpunctie).
    - Herkenning van meerdere namen in een tenaamstelling.
    - Geslachtsindicatie.
2) Validatie naam, voor:
    - Controleren van familienamen in combinatie met voegsels en adellijke titels.
    - Controleren op geslachtsconflicten.
3) Hoofdletterconversie, voor:
    - Kapitalisatie van persoonsnamen.
    - Kapitalisatie van bedrijfsnamen.

1.2 Kwaliteitsbepaling van adressen

Voor de kwaliteitsbepaling van adressen wordt de module HIquality Address van Human Inference ingezet. Hierbij wordt de functie 'Standaardiseren' gebruikt om een adres te controleren en eventueel te corrigeren (schrijffouten in bijvoorbeeld straatnaam, postcode of woonplaats).

Bij de controle van een adres worden de diverse componenten van een adres herkend, zoals:

1) Straatnaam.
2) Huisnummer.
3) Huisnummertoevoeging.
4) Postcode.
5) Woonplaats.

1.3 Kwaliteitsbepaling van telefoonnummers

HIquality Phone is een datakwaliteitscomponent om telefoonnummers te valideren. Daarbij wordt het volgende gedaan:

1) Bepalen van de juiste lengte.
2) Controleren van de landcode.
3) Checken van het netnummer en het abonneenummer.
4) Vaststellen van het type telefoonnummer, bijv. vast of mobiel nummer.
5) Formatteren van het telefoonnummer.
6) Ontleden van het telefoonnummer in logische delen.

1.4 Kwaliteitsbepaling van emailadressen

HIquality Email is een datakwaliteitscomponent om emailadressen te valideren. Daarbij wordt het volgende gedaan:

1) Bepalen van de syntax van de geadresseerde.
2) Bepalen van de syntax van het domein.
3) Controleren of het domein bestaat.
4) Ontleden van het emailadres in logische delen.

2 AGB-bestand

Dit hoofdstuk beschrijft de resultaten van de uitgevoerde schoning op de aangeleverde niet-verrijkte data van AGB. Het resultaat van de inhoudelijke validaties – dat is uitgevoerd met HIquality Name, HIquality Address, HIquality Phone en HIquality Email – is in het resultaat weergegeven met zogeheten traffic lights (verkeerslichten). Deze verkeerslichten hebben de volgende mogelijke waardes en betekenis:

- Green: De aangeboden gegevens zijn ongewijzigd door het betreffende datakwaliteitscomponent en daarmee correct.
- Green Corrected: Op de aangeboden gegevens zijn lichte wijzigingen doorgevoerd, zoals kleine typefoutjes of het gebruik van hoofdletters. Dergelijke correcties kunnen eenvoudig automatisch verwerkt worden.
- Orange: In de aangeboden gegevens zitten te veel fouten die aangepast moeten worden of ze zijn niet eenduidig, waardoor er meerdere suggesties zijn.
- Red: De aangeboden gegevens zijn foutief, een handmatige correctie is noodzakelijk.

2.1 Tenaamstelligen

2.1.1 Statistische gegevens

In totaal heeft Vektis 387.047 namen aangeleverd welke zijn onderverdeeld in zorgverleners (163.373), ondernemingen(115.725) en vestigingen (107.949).

In de onderstaande tabel is het ingedeeld in 4 hoofdgroepen met een afnemende mate van correctheid. Op basis van de correcties die HIquality Name al dan niet kan aanbrengen resulteert dit in 4 groepen zoals in de tabel gepresenteerd. Een naam wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt. Zo wordt een naam waarbij voornaam en achternaam omgedraaid zijn en de kapitalisatie aangepast is alleen opgenomen in de groep 'Word corrected; reliability is OK'.

Indien een tenaamstelling meerdere namen bevat wordt iedere individuele naam in de categorie 'Red' ingedeeld.

Beoordeling naam	Zorgverleners	Ondernemingen	Vestigingen
Green	0	318	319
Green Corrected	158674	107631	115406
Orange	991	0	0
Red	3708	0	0

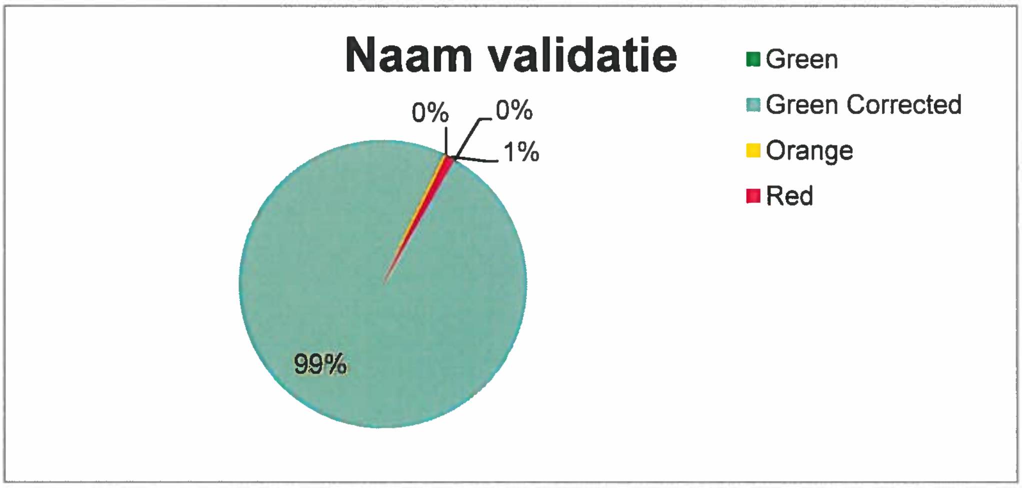
De onderstaande tabel met detailmeldingen is een nadere specificatie van de groepen uit de algemene statustabel. In deze tabel zijn de groepen met incorrecte namen als 3 hoofdcategorieën gepresenteerd. Ook bij deze tabel geldt dat een naam alleen in de laagste hoofdcategorie waarvoor die in aanmerking komt wordt opgenomen. Dit geldt tevens voor de onderverdeling in de kolom 'Manipulaties'.

Beoordeling naam	Manipulaties	Aantal	Percentage
Green	No problem in the name; reliability is OK ; no messages.	637	0.16
Green Corrected	Capitalization corrected; reliability is OK	381684	98.61
	Punctuation corrected; reliability is OK	26	0.01
	Word corrected; reliability is OK	1	0.00
Orange	Reliability is doubtful.	991	0.26
Red	Reliability is wrong. Problem with Family Name. Messages have been generated	1709	0.44
	Reliability is wrong. Problem with Title. Messages have been generated	11	0.00
	Reliability is wrong. Problem with Diacritics. Messages have been generated	121	0.03
	Reliability is wrong. Problem with Name Format. Messages have been generated	625	0.16
	Reliability is wrong. Problem with Prefix. Messages have been generated	732	0.19
	Reliability is wrong. Problem with Given Name. Messages have been generated	414	0.11
	Reliability is wrong. Problem with Gender. Messages have been generated	10	0.00

Beoordeling naam	Manipulaties	Aantal	Percentage
	Reliability is wrong. Missing Name. Messages have been generated	24	0.01
	Reliability is wrong. Problem with Name Element. Messages have been generated	60	0.02
	Reliability is wrong. Other problems. Messages have been generated	2	0.00

2.1.2 Kwaliteit van de tenaamstellingen

In onderstaande figuur is de kwaliteit van de tenaamstellingen grafisch weergegeven. De percentages worden gevormd door de aantallen uit de eerste tabel in deze paragraaf.



2.1.3 Voorbeelden

Id.	Naam	Beoordeling naam
1077443	H CORNELIS-V/D HEYDEN	Een punt in een voorvoegsel is verplicht: v/d""
1064955	M RAFSTAAKE	Onbekende achternaam: Rafstaake""
1083492	DAC PIERIE=SONNEVELD	Niet toewijsbaar deel in de naam: =""
1115734	HG BRUSSE-VANDEN NIEUWENDIJK	Achternaam bestaat niet met voorvoegsel vanden""

2.2 Adressen

2.2.1 Statistische gegevens

Alle adressen zijn beoordeeld op hun schrijfwijze conform de NEN-norm. In totaal zijn er 404.934 adressen beoordeeld.

In de onderstaande tabel zijn de adressen ingedeeld in 3 hoofdgroepen met een afnemende mate van correctheid. Op basis van de correcties die HIquality Address al dan niet kan aanbrengen resulteert dit in 4 groepen zoals in de tabel gepresenteerd.

Een adres wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt. Zo wordt een adres waarbij zowel de spelling van de straat als de kapitalisatie aangepast zijn alleen opgenomen in de groep 'adres is niet correct - eenduidige standaardisatie'.

Beoordeling adres	Bezoekadres	Correspondentieadres	Incassoadres	Tijdelijk verblijfadres	Verblijfadres / priveadres
Green	128165	271245	2	8	32
Green Corrected	1400	3291	0	0	0
Orange	59	127	0	0	7
Red	250	348	0	0	0

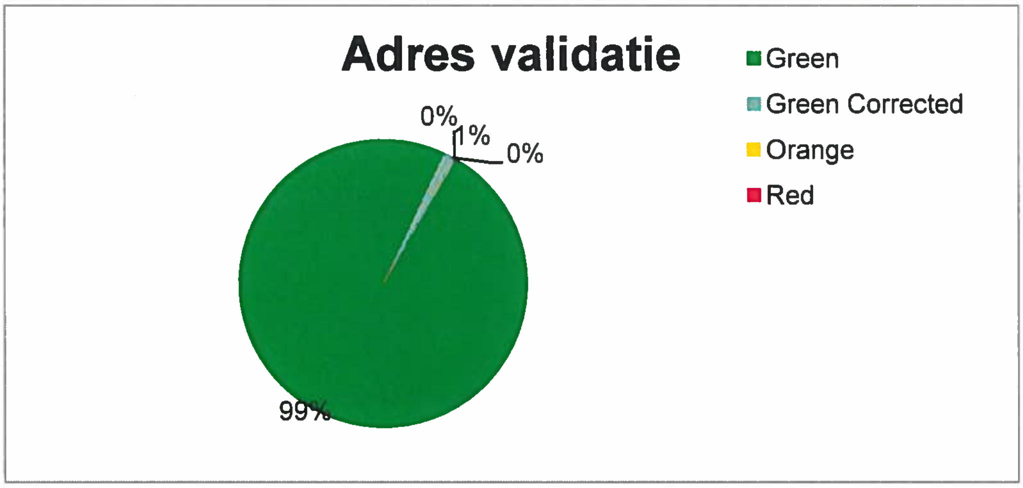
De onderstaande tabel met detailmeldingen is een nadere specificatie van de groepen uit de algemene statustabel. In deze tabel zijn de groepen met incorrecte adressen als 3 hoofdcategorieën gepresenteerd. Ook bij deze tabel geldt dat een adres alleen in de laagste hoofdcategorie waarvoor die in aanmerking komt wordt opgenomen.

Beoordeling adres	Manipulaties	Aantal	Percentage
Green	No problem in the address; reliability is okay ; no messages.	399452	98.65
Green Corrected	Capitalization corrected; reliability is okay	1412	0.35
	Diacritics corrected; reliability is okay	1	0.00
	Punctuation corrected; reliability is okay	11	0.00
	Word corrected; reliability is okay	509	0.13
	Other corrections; reliability is okay	2758	0.68
Orange	Reliability is doubtful.	193	0.05

Beoordeling adres	Manipulaties	Aantal	Percentage
Red	Reliability is wrong. Problem with thoroughfare. Messages have been generated	50	0.01
	Reliability is wrong. Problem with house number. Messages have been generated	45	0.01
	Reliability is wrong. Problem with postal code. Messages have been generated	30	0.01
	Reliability is wrong. Problem with posttown. Messages have been generated	1	0.00
	Reliability is wrong. Overall address is wrong. Messages have been generated	472	0.12

2.2.2 Kwaliteit van de adressen

Het figuur hieronder geeft een grafische weergave van de kwaliteit van de aangeleverde adressen. De percentages, zoals deze in de figuur gebruikt worden, komen uit de tabel “Algemene status over de adressen“ uit bovenstaande paragraaf.



2.2.3 Voorbeelden

Id.	Adresline1	Beoordeling Adres
104840	[redacted]	Adresveld toegevoegd. Veld(en) PostTown.
105020	[redacted]	Adresregels verplaatst (bijv. volgorde gewijzigd).
105029	[redacted]	Adresveld gecorrigeerd. Veld(en) PostTown.

2.3 Emailadressen

2.3.1 Statistische gegevens

Alle 167.280 emailadressen zijn beoordeeld op formaat, syntax en correctheid van de opgenomen domeinnaam. Tevens wordt, indien mogelijk, bij niet-correcte emailadressen een suggestie gedaan voor een correcte domeinnaam.

In de onderstaande tabel zijn de emailadressen ingedeeld in 4 hoofdgroepen met een afnemende mate van correctheid. Een emailadres wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt.

Indien de invoer uit meerdere emailadressen bestaat, wordt enkel het eerste emailadres beoordeeld en de rest genegeerd.

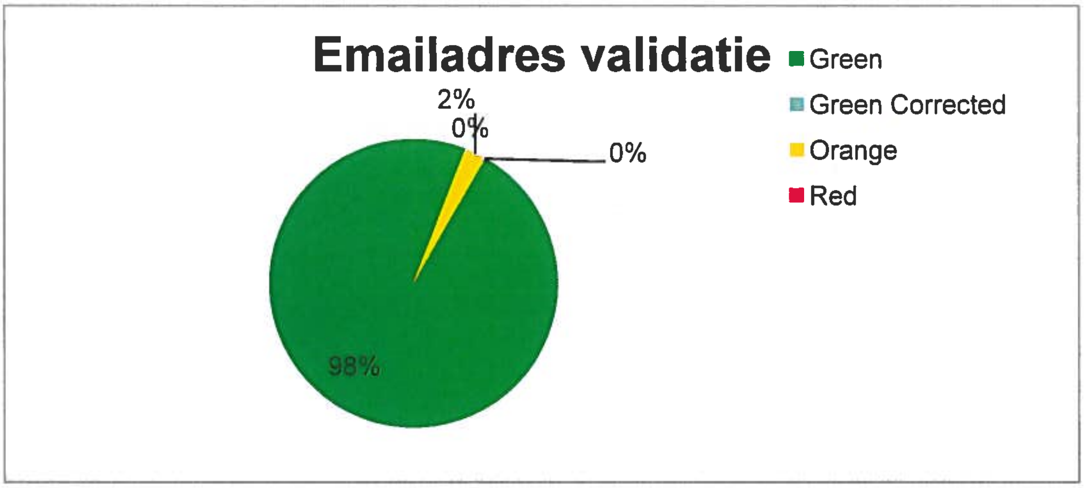
Beoordeling emailadres	Manipulaties	Aantal adressen gecontroleerd	Percentage
Green	Format is correct and the domain exists.	163177	97.55
Orange	Format is correct, but the domain is doubtful. The cause is either: The domain probably does not exist or does not respond to the lookup request.	4088	2.44
Red	Format is incorrect. Note that the domain check is not executed if the format is incorrect.	15	0.01

De onderstaande tabel met detailmeldingen is een nadere specificatie van de groepen uit de algemene statustabel.

Beoordeling emailadres	**Manipulaties**	**Aantal adressen gecontroleerd**	**Percentage**
Green	Valid domain	163177	97.55
Orange	Wrong or unknown domain	4088	2.44
Red	Invalid syntax of email	15	0.01

2.3.2 Kwaliteit van de emailadressen

In onderstaande figuur is de kwaliteit van de emailadressen grafisch weergegeven. De percentages komen uit de eerste tabel in deze paragraaf.



2.3.3 Voorbeelden

Id.	**Emailadres**	**Beoordeling emailadres**
297346	[redacted]	Valid domain
9	[redacted]	Wrong or unknown domain
394	[redacted]	Invalid syntax of email

2.4 Telefoonnummers

2.4.1 Statistische gegevens

Alle 362.214 telefoonnummers en 33.374 faxnummers zijn beoordeeld op formaat, lengte en geldigheid.

In de twee onderstaande tabellen zijn achtereenvolgens de vaste en mobiele nummers ingedeeld in 4 hoofdgroepen met een afnemende mate van correctheid voor. Een nummer wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt.

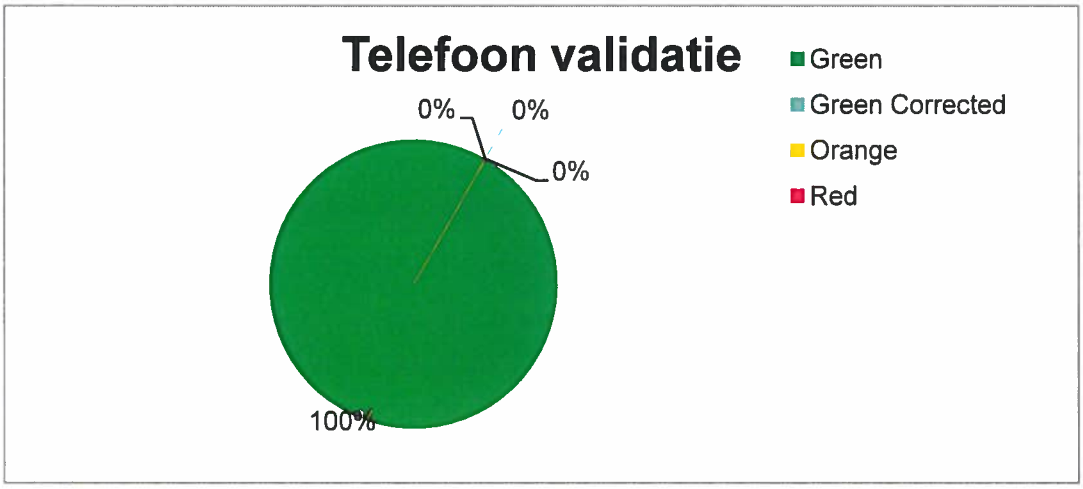
Beoordeling nummer	Telefoonnummer	Fax
Green	3327889	33317
Green Corrected	0	0
Orange	758	45
Red	193	12

De onderstaande tabel met detailmeldingen zijn een nadere specificatie van de groepen uit de algemene statustabel.

Beoordeling nummer	Manipulaties	Aantal nummers gecontroleerd	Percentage
Green	Geen Manipulaties	3360187	99.97
	Telefoonnummer hoort niet bij opgegeven land code	11	0.00
Orange	Telefoonnummer is twijfelachtig	803	0.02
Red	Telefoonnummer is geen nummer	14	0.00
	Telefoonnummer is te lang	57	0.00
	Telefoonnummer is te kort	134	0.00

2.4.2 Kwaliteit van de telefoonnummers

In onderstaande figuur is de kwaliteit van de telefoonnummers grafisch weergegeven. De percentages komen uit de eerste tabel in deze paragraaf.



2.4.3 Voorbeelden

Id	Telefoonnummer	Beoordeling telefoonnummer
111871	013-5079120	Telefoonnummer is twijfelachtig
116168		Telefoonnummer is geen nummer
220185	59997375287	Telefoonnummer hoort niet bij opgegeven land code
16579	0591-354744	Telefoonnummer is te lang
213294	00-32476523	Telefoonnummer is te kort

3 BIG-bestand

Dit hoofdstuk beschrijft de resultaten van de uitgevoerde schoning op de aangeleverde niet-verrijkte data van BIG. Het resultaat van de inhoudelijke validaties – dat is uitgevoerd met HIquality Address – is in het resultaat weergegeven met zogeheten traffic lights (verkeerslichten). Deze verkeerslichten hebben de volgende mogelijke waardes en betekenis:

- Green: De aangeboden gegevens zijn ongewijzigd door het betreffende datakwaliteitscomponent en daarmee correct.
- Green Corrected: Op de aangeboden gegevens zijn lichte wijzigingen doorgevoerd, zoals kleine typefoutjes of het gebruik van hoofdletters. Dergelijke correcties kunnen eenvoudig automatisch verwerkt worden.
- Orange: In de aangeboden gegevens zitten te veel fouten die aangepast moeten worden of ze zijn niet eenduidig, waardoor er meerdere suggesties zijn.
- Red: De aangeboden gegevens zijn foutief, een handmatige correctie is noodzakelijk.

3.1 Adressen

3.1.1 Statistische gegevens

Alle adressen zijn beoordeeld op hun schrijfwijze conform de NEN-norm (op basis van de BAG). In totaal zijn er 227.660 adressen beoordeeld.

In de onderstaande tabel zijn de adressen ingedeeld in 4 hoofdgroepen met een afnemende mate van correctheid. Op basis van de correcties die HIquality Address al dan niet kan aanbrengen resulteert dit in 4 groepen zoals in de tabel gepresenteerd. Een adres wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt. Zo wordt een adres waarbij zowel de spelling van de straat als de kapitalisatie aangepast zijn alleen opgenomen in de groep 'Reliability is wrong. Problem with thoroughfare. Messages have been generated'.

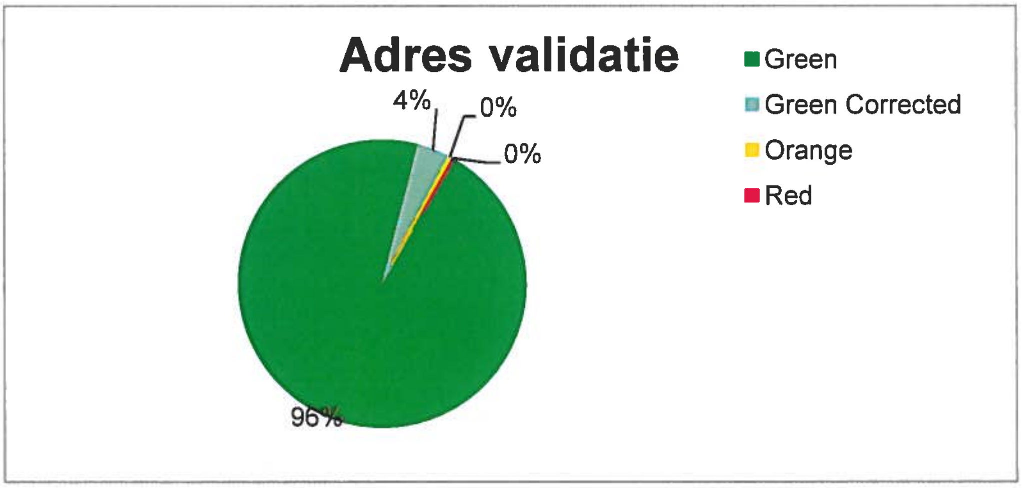
Beoordeling adres	Werkadres1	Werkadres2	Werkadres3
Green	202976	13290	1895
Green Corrected	7990	160	2
Orange	1147	21	0
Red	431	32	2

De onderstaande tabel met detailmeldingen is een nadere specificatie van de groepen uit de algemene statustabel. In deze tabel zijn de groepen met incorrecte adressen als 4 hoofdcategorieën gepresenteerd. Ook bij deze tabel geldt dat een adres alleen in de laagste hoofdcategorie waarvoor die in aanmerking komt wordt opgenomen.

Beoordeling adres	Manipulaties	Aantal	Percentage
Green	No problem in the address; reliability is okay ; no messages.	218161	95.71
Green Corrected	Capitalization corrected; reliability is okay	6481	2.84
	Word corrected; reliability is okay	1018	0.45
	Other corrections; reliability is okay	653	0.29
Orange	Reliability is doubtful .	1168	0.51
Red	Reliability is wrong. Problem with thoroughfare. Messages have been generated	6	0.00
	Reliability is wrong. Problem with house number. Messages have been generated	96	0.04
	Reliability is wrong. Problem with postal code. Messages have been generated	17	0.01
	Reliability is wrong. Problem with posttown. Messages have been generated	2	0.00
	Reliability is wrong. Overall address is wrong. Messages have been generated	344	0.15

3.1.2 Kwaliteit van de adressen

Het figuur hieronder geeft een grafische weergave van de kwaliteit van de aangeleverde adressen. De percentages, zoals deze in de figuur gebruikt worden, komen uit de tabel "Algemene status over de adressen" uit bovenstaande paragraaf.



Voorbeelden

Id.	**Address**	**Beoordeling Adres**
101491	[redacted]	Adresveld toegevoegd. Veld(en) PostTown.
102655	[redacted]	Adresveld gecorrigeerd. Veld(en) Thoroughfare.

4 Handelsregister-bestand

Dit hoofdstuk beschrijft de resultaten van de uitgevoerde schoning op de aangeleverde niet-verrijkte data van het Handelsregister. Het resultaat van de inhoudelijke validaties – dat is uitgevoerd met HIquality Address en HIquality Phone – is in het resultaat weergegeven met zogeheten traffic lights (verkeerslichten). Deze verkeerslichten hebben de volgende mogelijke waardes en betekenis:

- Green: De aangeboden gegevens zijn ongewijzigd door het betreffende datakwaliteitscomponent en daarmee correct.
- Green Corrected: Op de aangeboden gegevens zijn lichte wijzigingen doorgevoerd, zoals kleine typefoutjes of het gebruik van hoofdletters. Dergelijke correcties kunnen eenvoudig automatisch verwerkt worden.
- Orange: In de aangeboden gegevens zitten te veel fouten die aangepast moeten worden of ze zijn niet eenduidig, waardoor er meerdere suggesties zijn.
- Red: De aangeboden gegevens zijn foutief, een handmatige correctie is noodzakelijk.

4.1 Adressen

4.1.1 Statistische gegevens

Alle adressen zijn beoordeeld op hun schrijfwijze conform de NEN-norm. In totaal zijn er 466.513 adressen beoordeeld.

In de onderstaande tabel zijn de adressen ingedeeld in 4 hoofdgroepen met een afnemende mate van correctheid. Op basis van de correcties die HIquality Address al dan niet kan aanbrengen resulteert dit in 4 groepen zoals in de tabel gepresenteerd. Een adres wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt. Zo wordt een adres waarbij zowel de spelling van de straat als de kapitalisatie aangepast zijn alleen opgenomen in de groep 'Reliability is wrong. Problem with thoroughfare. Messages have been generated'.

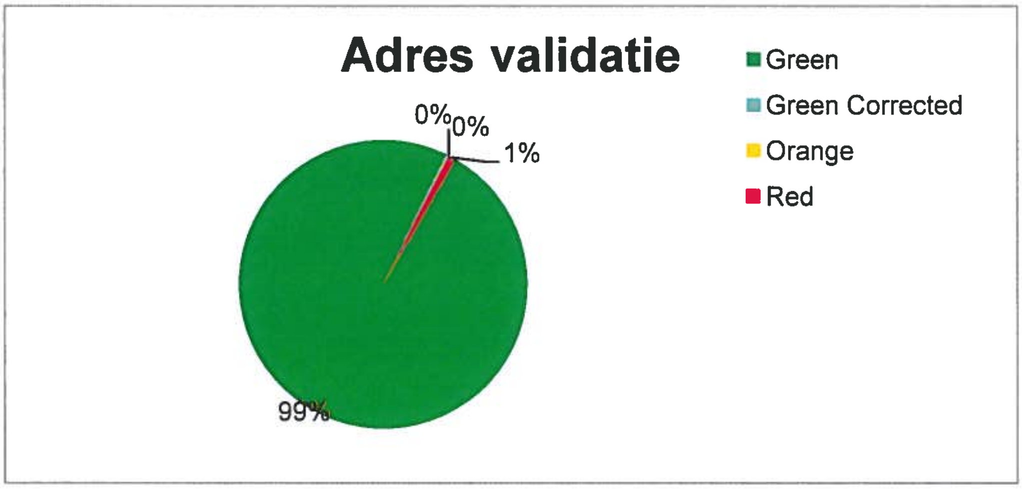
Beoordeling adres	Correspondentie adres	Vestigings adres
Green	222553	238626
Green Corrected	857	815
Orange	53	90
Red	1807	1712

De onderstaande tabel met detailmeldingen is een nadere specificatie van de groepen uit de algemene statustabel. In deze tabel zijn de groepen met incorrecte adressen als 4 hoofdcategorieën gepresenteerd. Ook bij deze tabel geldt dat een adres alleen in de laagste hoofdcategorie waarvoor die in aanmerking komt wordt opgenomen.

Beoordeling adres	Manipulaties	Aantal	Percentage
Green	No problem in the address; reliability is okay ; no messages.	461179	98.86
Green Corrected	Capitalization corrected; reliability is okay	115	0.02
	Word corrected; reliability is okay	356	0.08
	Other corrections; reliability is okay	1201	0.26
Orange	Reliability is doubtful .	143	0.03
Red	Reliability is wrong. Problem with thoroughfare. Messages have been generated	8	0.00
	Reliability is wrong. Problem with house number. Messages have been generated	1216	0.26
	Reliability is wrong. Problem with postal code. Messages have been generated	8	0.00
	Reliability is wrong. Problem with posttown. Messages have been generated	8	0.00
	Reliability is wrong. Overall address is wrong. Messages have been generated	2279	0.49

4.1.2 Kwaliteit van de adressen

Het figuur hieronder geeft een grafische weergave van de kwaliteit van de aangeleverde adressen. De percentages, zoals deze in de figuur gebruikt worden, komen uit de tabel "Algemene status over de adressen" uit bovenstaande paragraaf.



4.1.3 Voorbeelden

Id.	Adres	Beoordeling Adres
01080778	[redacted]	Gehele adres in de invoer is buitenlands/onjuist.
01100411	[redacted]	Gehele adres in de invoer is buitenlands/onjuist.
01117278	[redacted]	Gehele adres in de invoer is buitenlands/onjuist.

Zoals te zien is gaat het om dit geval om buitenlandse addressen. Echter in het aangeleverde bestand wordt geen landcode meegeleverd. Indien deze wel wordt meegeleverd, zouden deze addressen ook gecorrigeerd kunnen worden.

4.2 Telefoonnummers

4.2.1 Statistische gegevens

Alle 114.638 telefoonnummers en 96.942 mobiele nummers zijn beoordeeld op formaat, lengte en geldigheid.

In de twee onderstaande tabellen zijn achtereenvolgens de vaste en mobiele nummers ingedeeld in 4 hoofdgroepen met een afnemende mate van correctheid voor. Een nummer wordt ingedeeld in de laagste categorie waarvoor die in aanmerking komt.

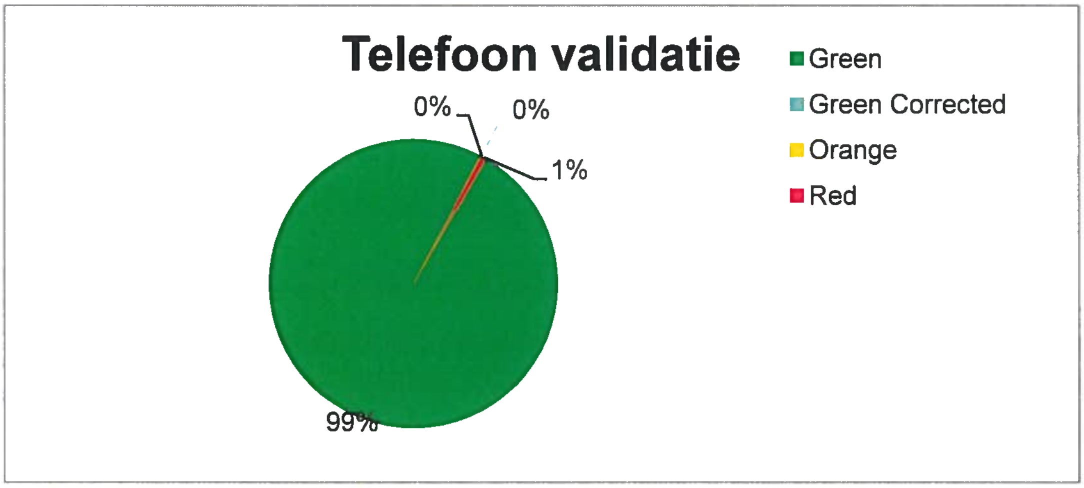
Beoordeling nummer	Telefoonnummer	Mobiel nummer
Green	114225	96159
Green Corrected	0	0
Orange	34	231
Red	379	552

De onderstaande tabel met detailmeldingen zijn een nadere specificatie van de groepen uit de algemene statustabel.

Beoordeling nummer	Manipulaties	Aantal nummers gecontroleerd	Percentage
Green	Geen Manipulaties	209587	99.06
	Telefoonnummer hoort niet bij opgegeven land code	528	0.25
			0.00
Orange	Telefoonnummer is twijfelachtig	427	0.20
			0.00
Red	Telefoonnummer is geen nummer	264	0.12
	Telefoonnummer is te lang	14	0.01
	Telefoonnummer is te kort	760	0.36

4.2.2 Kwaliteit van de telefoonnummers

In onderstaande figuur is de kwaliteit van de telefoonnummers grafisch weergegeven. De percentages komen uit de eerste tabel in deze paragraaf.



4.2.3 Voorbeelden

Id	Telefoonnummer	Beoordeling telefoonnummer
111871	013-5079120	Telefoonnummer is twijfelachtig
116168		Telefoonnummer is geen nummer
220185	59997375287	Telefoonnummer hoort niet bij opgegeven land code
16579	0591-354744	Telefoonnummer is te lang
213294	00-32476523	Telefoonnummer is te kort

Bijlage 5 Over Human Inference

Human Inference is opgericht in 1986 met als uitgangspunt software te ontwikkelen die menselijk redeneren simuleert en daarmee zorgt voor het optimaliseren van de kwaliteit van relatiegegevens. Door de op kennis gebaseerde oplossingen kan de software gegevens correct interpreteren en voor u de juiste beslissingen nemen dan wel voorstellen.

Toegevoegde waarde

Human Inference biedt veel toegevoegde waarde bij het creëren en onderhouden van datakwaliteit en bij het succesvol migreren dan wel samen laten komen van data. Als geen ander kent Human Inference de culturen en gewoontes en heeft zij die kennis verwerkt in haar oplossingen. Ook aspecten van meertaligheid zijn in onze oplossingen verwerkt. Wij kunnen met recht zeggen dat onze software het werk eenvoudiger maakt door uw medewerkers te ondersteunen bij het interpreteren, valideren en standaardiseren van gegevens en in besluiten over het al dan niet gelijk zijn van relaties. Vanaf de oprichting zijn wij bij onze klanten betrokken geweest in projecten voor het optimaliseren van klantcontacten, veelal gericht op het genereren van een centraal klantbeeld gebaseerd op correcte gegevens. Wij mogen vele grote ondernemingen en overheidsinstanties tot onze trouwe klanten rekenen.

Klanten binnen de overheid

Binnen de overheid werken wij onder andere samen met de Belastingdienst, de Justitiële Informatiedienst (JustID), de Basisvoorziening Vreemdelingen (BVV), de Dienst Uitvoering Onderwijs (DUO), de Personenserver Politie en de Rijksdienst voor Ondernemend Nederland. Daarnaast ondersteunt onze software authentieke registraties zoals het Handelsregister (Kamer van Koophandel) en de Kadastrale registratie (Kadaster).

Onze implementaties zijn bij deze organisaties gericht op onder meer:

- het genereren van een centraal klantbeeld, gebaseerd op golden records vanuit meerdere bronsystemen;
- het genereren van een verwijsindex waardoor eenvoudig te herkennen is in welke systemen een relatie is opgenomen;
- matching/ontdubbeling van relaties bij de totstandkoning van een centraal klantbeeld;

- het valideren, standaardiseren en verrijken van namen, adressen, telefoonnummers;
- het succesvol migreren van gegevens van oude systemen naar een nieuwe applicatie met optimale datakwaliteit;
- het controleren en bewaken van datakwaliteit aan de bron bij invoer van gegevens (First Time Right).

Meerwaarde

Steeds meer organisaties erkennen het belang van goede verwerking van relatiegegevens. Of het nu gaat om ontdubbeling van bestanden, online zoeken, invoercontrole of het samenvoegen van doublures: het is de bedoeling relaties al dan niet als gelijken te identiceren. Om een betrouwbare en zinvolle uitspraak over de mate van overeenkomst te kunnen doen, is toegevoegde kennis vereist. Kennis over de relatiegegevens en de omgeving en de cultuur waarin zij voorkomen. Deze kennis moet op een consistente en intelligente manier worden verzameld en vastgelegd en vervolgens in de identificatiemethode worden geïncorporeerd. Dat is feitelijk de enige juiste manier van verwerking van relatiegegevens.

Bijlage 6 Definities

Voor een goed begrip van de gehanteerde termen "zorgonderneming", "zorgverlener" en "zorgvestiging" zijn hieronder de definities opgenomen die in het architectuurdocument voor het LRZa, in de wet kwaliteit, klachten en geschillen zorg (Wkkgz) en door de bronhouders van het handelsregister en het AGB worden gehanteerd.

Zorgonderneming	
Architectuurdocument	Een zorgonderneming is een organisatorische eenheid met een rechtsvorm met of zonder rechtspersoonlijkheid, die bedrijfsmatig zorg en/of daaraan gerelateerde diensten verleent en/of doet verlenen door één of meer natuurlijke personen en/of rechtspersonen en/of organisatorische verbanden. Herkomst begrip / ontleend aan - Wkkgz, artikel 1 (zorgaanbieder, instelling en solistisch werkende zorgverlener) - Handelsregister (maatschappelijke activiteit) - Projectdiscussie inzake inhoudelijke scope LRZa (zie paragraaf 8.2)
Wkkgz	Zorgaanbieder is een instelling dan wel een solistisch werkende zorgverlener
Handelsregister	Van een onderneming is sprake indien een voldoende zelfstandig optredende organisatorische eenheid van één of meer personen bestaat waarin door voldoende inbreng van arbeid of middelen, ten behoeve van derden diensten of goederen worden geleverd of werken tot stand worden gebracht met het oogmerk daarmee materieel voordeel te behalen.
AGB	Een tot een zelfstandig rechtssubject behorend geheel van persoonlijke, materiële en immateriële factoren, waarmede op duurzame wijze een economisch doel wordt nagestreefd.
Zorgvestiging	
Architectuurdocument	Een Zorgvestiging is één uniek adresseerbaar pand waar de Zorgonderneming eenduidig benoemde zorg aanbiedt. Herkomst begrip / ontleend aan: - Wkkgz, artikel 1 (accommodatie) - Handelsregister (vestiging) - AGB (vestiging) - Afstemming met afnemers (ZIN, IGZ en NZa)
Wkkgz	Niet genoemd in de wet.
Handelsregister	Een gebouw of complex van gebouwen waar duurzame uitoefening van de activiteiten van een onderneming of rechtspersoon plaatsvindt.
AGB	Een gebouw of complex van gebouwen van waaruit duurzame uitoefening van (zorg)activiteiten van een Onderneming plaatsvindt.

Zorgverlener	
Architectuurdocument	Een natuurlijke persoon die gekwalificeerd en/of erkend is om beroepsmatig zorg te verlenen. Herkomst begrip / ontleend aan: - Wet BIG - artikel 3 - Wkkgz - artikel 1 - AGB (zorgverlener)
Wkkgz	Een natuurlijke persoon die beroepsmatig zorg verleent
Handelsregister	n.v.t.
AGB	Een natuurlijke persoon die beroepsmatig zorg verleent. Een natuurlijke of organisatorische rechtspersoon die de daadwerkelijke zorg aan de verzekerde verleent.
Zorgaanbod	
Architectuurdocument	Zorgaanbod zijn de vormen van zorg die door een Zorgonderneming op een Zorgvestiging worden geleverd. Herkomst begrip / ontleend aan - Wkkgz (artikel 12) - AGB - Afstemming met afnemers (ZIN, IGZ en NZa)
Wkkgz	Niet genoemd in de wet.
Handelsregister	n.v.t.
AGB	n.v.t.